NEPS
National Educational Panel Study

Information on Competence Testing

NEPS Starting Cohort 1 — Newborns

Education From the Very Beginning

Wave 6: 5 years

LIfBi
**LEIBNIZ INSTITUTE FOR
EDUCATIONAL TRAJECTORIES**

Test information	
Test information	Five-year-old children were tested individually in their homes in the presence of the anchor person and the interviewer.
Test sequence	The competence measures were administered in the following sequence: 1. Vocabulary: listening comprehension at word level (Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test-Revision IV German Adaption*, PPVT-IV. Frankfurt, Germany: Pearson) 2. Scientific competence 3. Delay of gratification: executive control The children either used the tablet to solve the tasks on their own (vocabulary and scientific competence) or answered verbally (delay of gratification).
Test duration (excluding setup)	approx. 40 minutes
Information about the administered competence measures	

Construct	Number of tasks	Duration (approx.)	Mode of administration	Next assessment
Vocabulary: listening comprehension at word level	19 sets with 12 tasks each (max.), with a stopping rule	14 minutes	visual stimuli presented on a tablet; each task featured one word and four possible selections; administered on a tablet	Wave 8 (2019)
Scientific literacy	20 tasks (max.)	20 minutes	technology-based assessment; picture-based multiple choice and multiple true-false; administered on a tablet	Wave 8 (2019)
Delay of gratification: executive control	-	6 minutes	physical objects; administered on a tablet	Wave 8 (2019)

Preface

The development of the individual tests is based on framework concepts. They are overarching concepts on the basis of which education-relevant competences are to be shown consistently and coherently over the entire personal history. Therefore, the following framework concepts, which served as a basis for the development of the test tools to measure the above-mentioned constructs, are identical in the different studies.

In addition to the competence measures, which are coherently assessed across the lifespan, stage-specific measures are assessed at specific points in time at which these measures are especially meaningful (cf. Berendes, Weinert, Zimmermann, & Artelt, 2013[1]). Usually, these assessments are not repeated.

---

[1] Berendes, K., Weinert, S., Zimmermann, S., & Artelt, C. (2013). Assessing language indicators across the lifespan within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online/Journal für Bildungsforschung Online*, *5*(2), 15–49.

Vocabulary: listening comprehension at word level

Listening comprehension at word, sentence and text/discourse level as indicators of linguistic competence in German

The importance of linguistic competences for school learning and for explaining social disparities in school careers is largely undisputed.

In the NEPS, German linguistic competences are captured, on the one hand, via listening comprehension at word, sentence and text/discourse level and, on the other hand, (from 2nd primary school class onwards) via indicators of reading skills (reading competence [text comprehension], reading speed). In Starting Cohort 1 of the NEPS, from children aged 3 years, listening comprehension is solely captured at word level and later on, in primary school, via indicators of reading ability.

Listening comprehension at word level (receptive vocabulary)

Measures of receptive vocabulary are a good and internationally applicable indicator of language skills and abilities acquired by children and adults. In numerous large international panel studies, passive vocabulary is collected as the central and sometimes sole indicator of cumulatively acquired linguistic-cognitive abilities taking into consideration individual basic skills (e.g., working memory capacity, speed variables) and environmental stimuli. Examples of such studies are the Head Start Family and Child Experiences Survey – FACES (USA)[2], the National Longitudinal Survey of Children and Youth – NLCSY (Canada; among others Lipps & Yiptong-Avila, 1999), the British Cohort Study – BCS70 (e.g., Bynner, 2004) or the European Child Care and Education (ECCE) Study, which is conducted in Germany, Austria, Spain and Portugal (e.g., European Child Care and Education (ECCE) Study Group, 1997).

The internationally most frequently used instrument for assessing receptive vocabulary is the Peabody Picture Vocabulary Test (PPVT), which is available in various versions (Dunn, 1959; Dunn & Dunn, 1981, 1997, 2007). Generally, the PPVT can be used for a very large age range (from 2.5 years to late adulthood) and is both easy to carry out and to analyze. A German version of the PPVT-IV (Dunn & Dunn, 2007; German version by Lenhard, Lenhard, Segerer, & Suggate, 2015) was used in Starting Cohort 1.

In the NEPS, the test was administered via a tablet PC. The children's task was to select one picture out of four, matching it to the word presented as an auditory cue by the tablet.

According to the guidelines of the PPVT-IV, the level of difficulty (test entry and termination) varies depending on the age and performance of the children. The practice unit at the start of the test also varies depending on the age and performance of the children. Once a child has solved at least two tasks correctly during the practice phase, he or she moves on to the test phase. The test consists of a total of 19 sets with staggered levels of difficulty, each set consisting of 12 items.

Testing procedure in this wave: The test begins with a practice phase consisting of at least two and a maximum of six tasks. The starting set depends on performance in the practice phase as well as on the age of the children. If the child makes more than one mistake in the starting set, the next lower set follows until a maximum of one mistake is made in one set (basal set). The test is then carried out until the ceiling set – the set in which the child has made more than seven mistakes – has been identified; the sets that have already been processed are left out.

---

[2] http://www.acf.hhs.gov/programs/opre/hs/faces/

The Scientific Use File contains the number of administered practice items, the correctness of the answers for each test item (correct, false), the basal set and the ceiling set. In addition, the sum score is included, indicating the number of correctly solved items. All items that are in lower sets than the basal set are assumed to be correct.

References

Bynner, J. (2004). Participation and progression: use of British Cohort Study data in illuminating the role of basic skills and other factors. *Nuffield Review of 14-19 Education and Training*, p. Working Paper 9.

Dunn, L. M. (1959). *Peabody Picture Vocabulary Test (PPVT): Manual of directions and forms.* Nashville, TN: American Guidance Service.

Dunn, L. M. & Dunn, D. M. (1981). *Peabody Picture Vocabulary Test-Revised (PPVT-R).* Circle Pines, MN: American Guidance Service.

Dunn, L. M. & Dunn, D. M. (1997). *Peabody Picture Vocabulary Test, Third Edition (PPVT-III).* Circle Pines, MN: American Guidance Service.

Dunn, L. M. & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4).* Upper Saddle River, NJ: Pearson.

European Child Care and Education (ECCE) Study Group. (1997). *European Child Care and Education.* Berlin, Germany: Freie Universität Berlin, Department of Education.

Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test-4.* Frankfurt, Germany: Pearson.

Lipps, G. & Yiptong-Avila, J. (1999). From home to school: how Canadian children cope. *Education Quarterly Review, 6*(2), 51–57.

Scientific literacy

Scientific literacy is the precondition for participating in world affairs marked by science and technology (Prenzel, 2000; Prenzel et al., 2001; Rost et al., 2004) and is viewed as a predictor for an economically, socially and culturally successful life. Many problems and issues we encounter in our daily life require an understanding of natural sciences and technology. Scientific topics and problems affect all people. Therefore, current discussions on the goals of scientific education focus on the concept of scientific literacy for all people (Osborne & Dillon, 2008). Such literacy is the basis for lifelong learning, serves as a connection for further learning (OECD, 2006; Prenzel et al., 2007) and, thus, also influences professional careers.

Based on this, the NEPS definition of scientific literacy follows the Anglo-Saxon literacy concept (Bybee, 1997; Gräber, Nentwig, Koballa & Evans, 2002; OECD, 2006) which does not regard scientific competence as the simple reproduction of acquired knowledge but rather as the flexible use of acquired knowledge in different situations and contexts of daily life.

In the NEPS, scientific literacy is understood as the use of scientific knowledge in the environmental, technological and health contexts (Hahn et al., 2013). In addition, the concept distinguishes between content-related and process-related elements (see Fig. 1). Knowledge of science comprises content-related *matter*, *systems*, *development* and *interactions*. Knowledge about science includes *enquiry and scientific reasoning* which involve, among other things, checking hypotheses, interpreting findings as well as measuring principles and measuring error control.
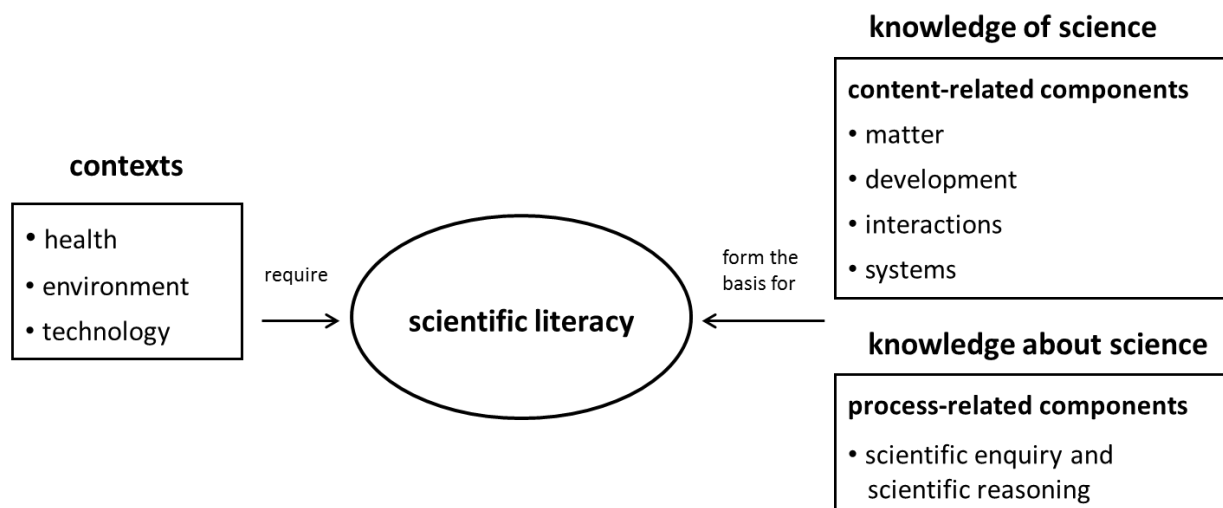
knowledge of science

content-related components
- matter
- development
- interactions
- systems

contexts
- health
- environment
- technology

require → scientific literacy ← form the basis for

knowledge about science

process-related components
- scientific enquiry and scientific reasoning

Fig.1. Application contexts as well as content-related and process-related elements of the NEPS scientific literacy test (Hahn et al., 2013).

To select its contexts as well as the content-related and process-related elements, the NEPS uses PISA (OECD, 2006), the *Benchmarks for Scientific Literacy* of the *American Association for the Advancement of Science* (AAAS, 2009) and the education standards of the Conference of Ministers of Education for the medium-level school-leaving qualification (KMK, 2005a, 2005b, 2005c) as a guideline. The selected contexts are of personal, social and global relevance. Considering current scientific research and the general events of the day, it is assumed that they will remain important across the entire life span of the test persons. Figure 2 provides an overview of the overlap of content-related components between PISA, the German educational standards and the NEPS. The selected content-related and process-related elements cover central concepts of all scientific disciplines.
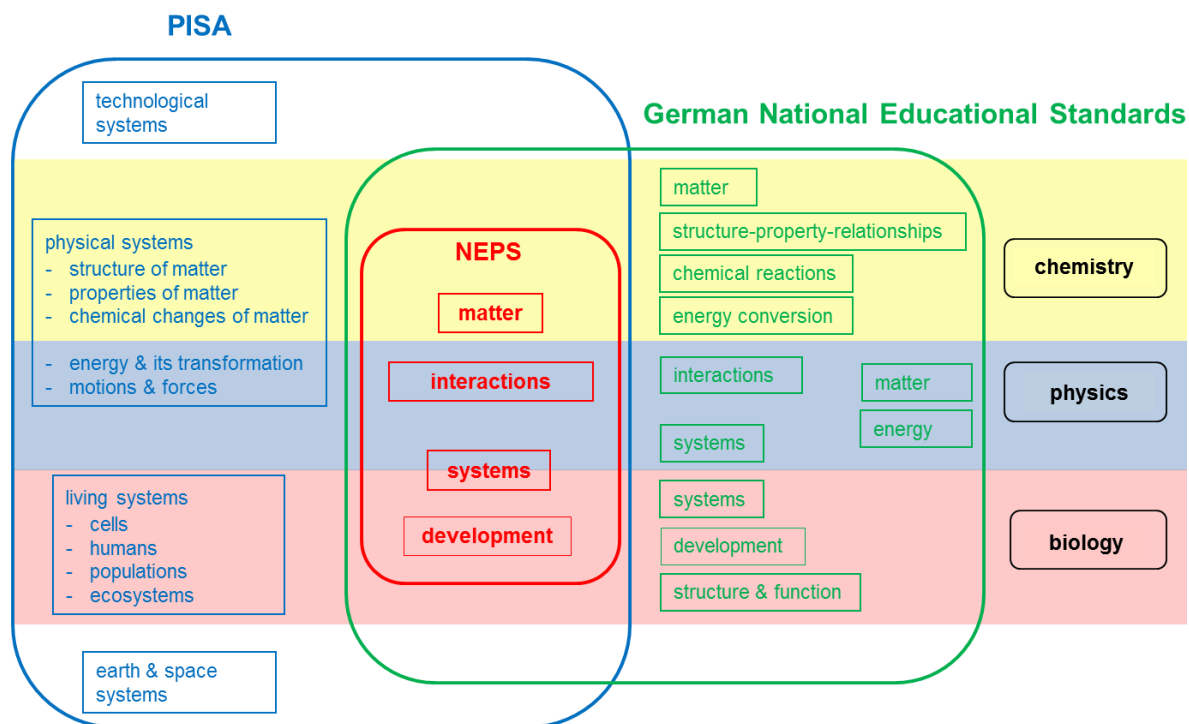
Fig.2. Overview of the overlap of content-related components between PISA, the German educational standards and the NEPS (Hahn et al., 2013).

The scientific literacy of 5-year-old children in Starting Cohort 1 of the NEPS is assessed using a tablet-based test. The test items are embedded in a "science and technology game", and a little dragon called "Nepsi" guides the children through the game. He reads the picture-based items and possible answers to them and then asks the children either to pick the right answer out of four pictures (multiple choice) or to judge whether successively shown pictures present a right or wrong answer (multiple true-false).

In the end, one scientific literacy score is computed and published in the Scientific Use File for Starting Cohort 1.

References

American Association for the Advancement of Science. (1989). *Science for all Americans: A Project 2061 Report on goals in science, mathematics and technology*. Washington, DC: American Association for the Advancement of Science.

American Association for the Advancement of Science. (AAAS). (2009). *Benchmarks for science literacy*. *Project 2061*. Retrieved from http://www.project2061.org/publications/bsl/online/index.php

Bybee, R. W. (1997). Towards an understanding of scientific literacy. In W. Gräber & C. Bolte (Eds.). *Scientific literacy – An international symposium* (pp. 37–68). Kiel, Germany: Institute for Science and Mathematics Education (IPN).

Gräber, W., Nentwig, P., Koballa, T. & Evans, R. (Eds.). (2002). *Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung*. Opladen: Leske + Budrich.

Hahn, I., Schöps, K., Rönnebeck, S., Martensen, M., Hansen, S., Saß, S., Dalehefte, I. M., & Prenzel, M. (2013). Assessing scientific literacy over the lifespan – A description of the NEPS science framework and the test development. *Journal for Educational Research Online*, *5*(2), 110–138.

KMK (2005a). *Resolutions of the Conference of Ministers of Education: Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Resolution of 16.12.2004. Munich, Germany: Luchterhand.

KMK (2005b). *Resolutions of the Conference of Ministers of Education: Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. Resolution of 16.12.2004. Munich, Germany: Luchterhand.

KMK (2005c). Resolutions of the Conference of Ministers of Education*: Bildungsstandards im Fach Physik für den Mittleren Schulabschluss.* Resolution of 16.12.2004. Munich, Germany: Luchterhand

OECD (2006). *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006*. Paris: Organisation for Economic Co-Operation and Development.

Osborne, J. & Dillon, J. (2008). *Science education in Europe: Critical reflections. A report to The Nuffield Foundation*. London, United Kingdom: King's College.

Prenzel, M. (2000). Lernen über die Lebensspanne aus einer domänenspezifischen Perspektive: Naturwissenschaften als Beispiel. In F. Achtenhagen & W. Lempert (Eds.), *Lebenslanges Lernen im Beruf - seine Grundlegung im Kindes- und Jugendalter. Volume IV. Formen und Inhalte von Lernprozessen* (pp. 175–192). Opladen, Germany: Leske + Budrich.

Prenzel, M., Rost, J., Senkbeil, M., Häußler, P., & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Eds.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 191–248). Opladen, Germany: Leske + Budrich.

Prenzel, M., Schöps, K., Rönnebeck, S., Senkbeil, M., Walter, O., Carstensen, C. H., & Hammann, M. (2007). Naturwissenschaftliche Kompetenz im internationalen Vergleich. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Eds.), *PISA 2006 – Die Ergebnisse der dritten internationalen Vergleichsstudie* (pp. 63–105). Münster, Germany: Waxmann.

Rost, J., Prenzel, M., Carstensen, C.-H., Senkbeil, M., & Groß, K. (Ed.). (2004). *Naturwissenschaftliche Bildung in Deutschland. Methoden und Ergebnisse von PISA 2000*. Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Delay of gratification: executive control

Self-regulation is defined as the ability to control and manage one's own thinking, feeling and actions (Neubauer, Gawrilow, & Hasselhorn, 2011) as well as to plan, pursue and consequently achieve personal goals (Zimmermann, 2000). Self-regulation abilities include different and diverse facets, which are measured accordingly (e.g., self-/external judgement, direct standardized observations, experimental tasks). The NEPS covers different aspects and facets of self-regulation, aspects of

cognitive self-regulation in the form of procedural and declarative metacognition (Weinert et al., 2019) as well as various survey items designed to capture emotional and behavioral self-control.

The NEPS Starting Cohort 1 (Newborns) and Starting Cohort 2 (Kindergarten) also use delay of gratification[3] tasks (for Starting Cohort 2, see Luplow, Schönmoser, Lorenz, & Schmitt, 2019). To assess delay of gratification, two procedures are implemented: the waiting paradigm and the choice paradigm (Mischel, 1974; Mischel, 2015). The waiting paradigm, which is used in Starting Cohort 1 for children aged 3 and 5 years, is used to examine the child's ability to delay an immediate small reward in favor of a bigger time-delayed reward (Mischel & Gilligan, 1964). Thus, the paradigm constitutes a reaction inhibition task (Garon, Bryson, & Smith, 2008), which can conceptually be assigned to executive functions, which are significant for controlling and planning actions and behavior (Neubauer et al., 2011). In Starting Cohort 1, executive functions are also measured at the age of 4 years with a flanker task.

The ability to self-regulate and cognitively process different types of behavior (Mischel, 1974) is considered to be important for child development. A number of studies indicate that well-developed abilities of self-regulation at preschool age are predictive of later academic performance, the ability to cope with stress, the development of socio-emotional competencies and the ability to concentrate (Baumeister & Vohs, 2004; Kochanska, Murray, & Coy, 1997; Tangney, Baumeister, & Boone, 2004; Wulfert, Block, Ana, Rodriquez, & Colsman, 2002), although these findings have also been the subject of criticism (Watts, Duncan, & Quan, 2018). Furthermore, there is evidence that the decision in favor of an immediate small gift or a bigger gift at a later time is related to the age of children, social background characteristics and other child measures (Lemmon & Moore, 2007; Thompson, Barresi & Moore, 1997; Watts et al., 2018). The tendency to wait usually rises with increasing age (Lorenz, Schmitt, Luplow, & Schönmoser, 2016).

In this wave, both the setting and the waiting conditions (Mischel, 1974) were standardized: two wrapped gifts – one big and one small – were presented to the child, and a USB button was placed between the gifts. The child could either wait for a certain amount of time, which was unknown to the child, to receive the big gift or press the button to receive the small gift immediately. Hence, this decision determined which gift was selected and the waiting time. The maximum waiting time was 5 minutes (301 seconds), after which the child automatically received the big gift.

In addition, the behavior of the child during the waiting time was observed by the interviewer via a time-sampling method. The behavior observation categories were entered on the tablet computer. There were ten intervals of 30 seconds each for entering the respective behavior category. The interviewer could choose between the following types of behavior: (the child) gets up and leaves; looks at the gift; looks at the button; talks; (sits) restlessly on his/her chair; does other things. In addition to information about the observed behavior, the intervals after the child decided to end the test were coded as "not reached".

The following variables are included in the Scientific Use File: the waiting time (in seconds); the child's decision (big vs. small gift); all behavior intervals.

---

[3] Note: In the NEPS, the terms "delay of gratification" and "delayed gratification" are used synonymously.

References

Baumeister, R. F. & Vohs, K. D. (2004). *Handbook of self-regulation: Research, theory, and applications.* New York: The Guilford Press.

Garon, N., Bryson S. E., & Smith I. M. (2008). Executive Function in Preschoolers: A Review Using an Integrative Framework. *Psychological Bulletin*, *134*(1), 31–60.

Kochanska, G., Murray, K. T., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*(2), 263–277.

Lemmon, K., & Moore, C. (2007). The development of prudence in the face of varying future rewards. *Developmental Science, 10*(4), 502–511.

Lorenz, C., Schmitt, M., Luplow, N., & Schönmoser, C. (2016). Soziale Disparitäten im Vorschulalter und der Einfluss der Selbstregulation. *Zeitschrift für Grundschulforschung, 9*(1), 65–77.

Luplow, N., Schönmoser, C., Lorenz, C., & Schmitt, M. (2019). *Die Messung des Belohnungsaufschubes in der Startkohorte 2 des Nationalen Bildungspanels (NEPS) im Kindergarten und der Grundschule (NEPS Survey Paper No. 54).* Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study (NEPS).

Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz, *Advances in Experimental Social Psychology* (Vol. 7, pp. 249–292). New York, NY: Academic Press.

Mischel, W. (2015). *Der Marshmallow Test.* Munich, Germany: Siedler Verlag.

Mischel, W. & Gilligan, C. (1964). Delay of gratification, motivation for the prohibited gratification, and responses to temptation. *The Journal of Abnormal and Social Psychology, 69*(4), 411–417.

Neubauer, A., Gawrilow, C., & Hasselhorn, M. (2011). Belohnungsaufschub: Ein Ansatz zur Frühprognose volitionaler Kompetenzen. In M. Hasselhorn & W. Schneider (Eds.), *Frühprognose schulischer Kompetenzen* (pp. 202–220). Göttingen, Germany: Hogrefe.

Tangney, J. P., Baumeister, R. F., & Boone, A. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271–324.

Thompson, C., Barresi, J., & Moore, C. (1997). The development of future-oriented prudence and altruism in preschoolers. *Cognitive Development, 12*(2), 199–212.

Watts, T. W., Duncan, G. J., & Quan, H. (2018). Revisiting the Marshmallow Test: A Conceptual Replication Investigating Links Between Early Delay of Gratification and Later Outcomes. *Psychological Science, 29*(7), 1159–1177.

Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T., Carstensen, C. H., & Lockl, K. (2019). Development of Competencies Across the Life Course. In H.-P. Blossfeld, & H.-G. Roßbach (Eds.), *Education as a Lifelong Process: The German National Educational Panel Study (NEPS)* (pp. 57–82). Wiesbaden, Germany: Springer VS.

Wulfert, E., Block, J. A., Ana, E. S., Rodriguez, M. L., & Colsman, M. (2002). Delay of Gratification: Impulsive Choices and Problem Behaviors in Early and Late Adolescence. *Journal of Personality, 70*(4), 533–552.

Zimmermann, B. J. (2000). Attaining Self-Regulation: A Social Cognitive Perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). San Diego, CA: Academic Press.