Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Sciences de l'Éducation

Arrêté ministériel : 25 mai 2016

Présentée par

Martin Galilée

Thèse dirigée par **Erica de Vries**, Professeur, UGA, et
Co-dirigée par **Kristine Lund**, Ingénieur de Recherche, HDR, ENS Lyon

préparée au sein du **Laboratoire de Recherche sur les Apprentissages en Contexte (LaRAC)**
dans **l'École Doctorale Sciences de l'Homme, du Politique et du Territoire (SHPT)**

Représentations externes pour l'apprentissage et la comparaison de la consommation d'énergie

External representations for learning and comparing energy consumption

Thèse soutenue publiquement le **14 décembre 2017**
devant le jury composé de :

Madame Erica de Vries

Professeur, Université Grenoble Alpes, Directrice de thèse

Madame Kristine Lund

Ingénieur de Recherche, HDR, ENS Lyon, Co-directrice de thèse

Madame Shaaron Ainsworth

Professeur, University of Nottingham, Rapporteur

Monsieur Gérard Sensevy

Professeur, ESPÉ de Bretagne, Rapporteur

Monsieur Jean-Michel Boucheix

Professeur, Université de Bourgogne, Examinateur

Madame Karine Mazens

Maître de Conférences, Université Grenoble Alpes, Examinateur

ABSTRACT

In this thesis is first considered how energy is taught and learned about in school, focusing on the discrepancies between a scientific definition of energy and a societal definition of energy, and discussing units of energy and the confusion they induce. Perspectives for education and energy management are provided. Then, focus is placed on the representations of energy provided in home energy management systems, seeking to propose an original classification based on educational strategies. The major obstacles met by designers reveal how energy management tools can be adapted to human cognition. Next, human numerical and magnitude processing abilities are discussed in depth, taking the viewpoint of grounded cognition and building a framework through which the impact of external representations of energy on learning and comparing can be established, understood, and predicted. This leads to two empirical studies.

The first study tests the effect of external representation (symbolic or spatial) on recall and comparisons from memory. Accuracy and response time at comparisons are used as dependent variables. Results indicate analog processing of magnitude in both conditions, and show that external representation affects performance at both recall and comparison, with symbolic external representation increasing recall and comparison accuracy, and spatial external representation increasing comparison speed. The second study tests the effects of spatiality, groundedness, and physicality in external representations, also on recall and comparisons from memory, using the same dependent variables. Results indicate analog processing in all conditions. Spatiality decreases recall accuracy but increases comparison speed. Groundedness and physicality show no effect.

Results are consistent with grounded cognition's mental simulations hypothesis (Barsalou, 1999, 2008; Wilson, 2002) as well as Dehaene's (1997) view on numerical cognition, in which number sense is based on a continuous accumulator that does not directly process discrete numbers. Theoretical implications and practical applications are discussed.

RÉSUMÉ

Dans cette thèse est d'abord considéré comment l'énergie est enseignée et apprise à l'école, montrant les divergences entre définition scientifique et sociétale de l'énergie, et considérant les unités d'énergie et la confusion qu'elles engendrent. Des perspectives pour l'éducation et la gestion de l'énergie sont présentées. Ensuite, l'attention est portée sur les représentations de l'énergie proposées par les systèmes domestiques de gestion, et une classification originale basée sur des stratégiques didactiques est proposée. Les obstacles majeurs rencontrés par les designers révèlent comment les outils de gestion de l'énergie peuvent être adaptés à la cognition humaine. Enfin, les capacités humaines de traitement des grandeurs numériques sont examinées en profondeur du point de vue de la cognition incarnée. Un cadre est construit au travers duquel l'impact des représentations externes de l'énergie sur l'apprentissage et la comparaison peut être établi, compris, et prédit. Ceci mène à deux études empiriques.

La première étude teste l'effet de la représentation externe (symbolique ou spatiale) sur le rappel et la comparaison de mémoire. Précision et temps de réponse sont les variables dépendantes dans la comparaison. Les résultats indiquent un traitement analogique dans les deux conditions. La représentation externe symbolique accroît la précision dans le rappel et la comparaison, et la représentation externe spatiale accroît la vitesse de comparaison. La seconde étude teste l'effet de la spatialité, de l'ancrage, et de la physicalité dans les représentations externes, également sur le rappel et les comparaisons de mémoire, utilisant les mêmes variables dépendantes. Les résultats indiquent un traitement analogique dans toutes les conditions. La spatialité décroît la précision dans le rappel mais accroît la vitesse de comparaison. Ancrage et physicalité n'ont pas d'effet.

Les résultats corroborent l'hypothèse de la cognition ancrée sur les simulations mentales (Barsalou, 1999, 2008; Wilson, 2002) ainsi que la perspective de Dehaene (1997) sur la cognition numérique, dans laquelle le sens du nombre est basé sur un accumulateur analogique et non discret. Implications théoriques et applications pratiques sont discutées.

ACKNOWLEDGEMENTS

This thesis bears the name of a single author, but he is himself the product of an ongoing collective work. I first want to thank my advisor Erica de Vries for her passion and her patience, and for driving this project from the first day until the last. I also want to thank all my colleagues in France or elsewhere. I think about everyone at the LaRAC (ex-LSE) in Grenoble, at Modoc Hall in Chico, California, and at the IWM Leibnitz-Institut für Wissenmedien in Tübingen, Germany. It was wonderful to be surrounded by such great minds and generous people. My gratitude especially goes to Neil H. Schwartz and Katharina Scheiter who put their trust in me and welcomed me into their labs. These were times. I also want to thank Kris Lund for her uninterrupted support and attention. I wish I spent more time working in Lyon. I also want to thank the institutions who made this thesis possible, including the Région Auvergne-Rhône-Alpes and ARC5 who funded me for three years. Thanks also to Paul for the financial support.

I also want to thank all those who supported me in the academic and professional endeavor that my personal life became, especially in the hard times of writing: Éloïse, Adrien, Félix, Lola, Kevin, Julia, Mélissa, and Maxime. Finally, thanks to Eduardo for all his support and the times we shared. I hope you soon find yourself writing similar lines.

CONTENTS

1

INTRODUCTION

World fossil fuel reserves are depleting, while the waste product of their combustion is altering the climate that saw mankind and modern civilization rise and spread. Society is presently at the dawn of an energy crisis. Citizens of today and tomorrow will need to make choices involving energy at every level, from the use of appliances to the election of national leaders. Power plants will be built and dismantled, investments will be made in research and technology, and ubiquitous energy producing sites will sprout all over landscapes and skylines. Today, citizens can purchase solar panels and heat pumps, with the option to sell energy to the grid. They can choose their home energy provider according to the energy source it uses. They also have a range of fuel options for their vehicles, and are even required to make an educated choice when purchasing home equipment. Three options of equivalent lightbulbs are thus on display on the shelves of every hardware store. The importance of decisions regarding energy and the options available will only keep widening as society provides innovative solutions to the energy crisis. Citizens will need to be prepared; they will need to know and learn about energy and its management.

Education professionals must be ready to teach more than a scientific concept of energy, but also a concept relevant to everyday life and societal scale. Energy providers must be ready to provide practical and understandable feedback to energy users. Designers must be ready to create intelligible visualizations tracking energy consumption. The present work seeks to provide ways to reach these goals by inquiring into the science of human learning and the mechanisms of cognition.

This thesis will begin with three theoretical chapters, follow with two empirical studies each in a chapter, and end with a general discussion. The first theoretical chapter will consider how energy is taught and learned about in school, focusing on the discrepancies between a scientific definition of energy and a societal definition of energy, and discussing the units of energy and the confusion they induce. The goal of the chapter is to portray the current reality of teaching and learning about energy in order to determine perspectives for education as well as energy management. The second chapter will focus on the representations of energy provided in home energy management systems, seeking to propose an original classification based on educational strategies. This review will provide insights on the major obstacles met by designers to provide cognitively adapted energy management tools. The third theoretical chapter will discuss in depth human numerical and magnitude processing abilities, taking the viewpoint of grounded cognition and establishing a framework through which the impact of external representations of energy on learning and comparing can be established, understood, and predicted. This chapter will lead to two empirical studies.

The first empirical study will consider the effects of learning about energy consumption of a set of appliances with different external representations in tasks of recall of energy consumption and comparison of energy consumption. A symbolic (digital) external representation and a spatial (graphical) external representation will be studied. The second empirical study will use a similar method, while broadening the range of external representations considered, in order to investigate the effects of spatiality, groundedness, and physicality of external representations. Findings of both studies will be examined in a final general discussion chapter.

2

LEARNING ABOUT ENERGY

The aim of this chapter is to identify the essential obstacles in learning about energy. These obstacles are related to gaps between the societal and the scientific definitions of the concept, to misconceptions arising from these gaps, and to confusing units. This chapter presents a review of the literature on teaching and learning about energy and a reflection on energy units.

2.1 Learning two definitions of energy

There are two different definitions of energy: the scientific definition and the societal definition. The scientific definition of energy is taught in school because school seeks to train future scientists, engineers, and technicians whose jobs will require mastery of scientific concepts such as energy. School also educates young citizens who need to take position on complex societal problems and decide the present and the future of society (Doménech et al., 2007). Citizens, including those who are science professionals, need to understand energy in its societal definition and properties. Consequently, the societal definition is taught in schools alongside the scientific definition (Doménech et al., 2007; Nordine, Krajcik, & Fortus, 2011; Vince & Tiberghien, 2012). Learners are thus confronted with two definitions, corresponding to domains (Vince & Tiberghien, 2012), or worlds of energy (Hervé, Venturini, & Albe, 2014). Several discrepancies set these definitions apart.

The first discrepancy between the two definitions lies in the nature of energy. As a non-scientific concept, the societal definition of energy does not have a clear definition. Loosely speaking, it is a quasi-material entity which allows putting mechanical, physical, or biological mechanisms into action, and which gets consumed when this happens. This definition is a vulgarization of the scientific definition, which is actually unclear itself.

Because the concept of energy is one of the most abstract ones in physics, new scientific definitions of energy are still proposed and debated in the present day (M. Bächtold, Munier, Guedj, Lerouge, & Ranquet, 2014). The scientific concept of energy was originally built in reference to the physical concept of work, the scientific definition of which also differs greatly from the everyday definition (Quinn, 2014; Watts, 1983), and does not provide a satisfying basis to define energy. Poincaré (1902) stated that it is impossible to give a general definition of energy, except by the vague statement: "there is something that remains constant" (Hervé et al., 2014). Richard Feynman taught and defined energy as a mere mathematical principle, in complete opposition with the societal view of energy as a quasi-material substance:

"There is a fact, or if you wish, a law, governing all natural phenomena that are known to date. There is no known exception to this law—it is exact so far as we know. The law is called the conservation of energy. *It states that there is a certain quantity, which we call energy, that does not change in the manifold changes which nature undergoes. That is a most abstract idea, because it is a mathematical principle; it says that there is a numerical quantity which does not change when something happens. It is not a description of a mechanism, or anything concrete; it is just a strange fact that we can calculate some number and when we finish watching nature go through her tricks and calculate the number again, it is the same." (Feynman, Leighton, & Sands, 1965, Chapter 4)*

Nonetheless, energy is sometimes also described as a quasi-material substance in science education. Watts (1983) describes various frameworks by which energy can be understood. One of them, very prevailing, likens energy to a fluid. Within this framework, energy is quasi-material: it flows and can be transported, stored, and passed on. Different frameworks can be thought of as metaphors, which according to Lakoff and Johnson (1980) allow the acquisition of new abstract concepts by linking previously acquired concepts to new ones. This scaffolding by metaphors is ultimately founded on the bedrock of direct physical experience of the world. Some authors (Colonnese, Heron, Michelini, Santi, & Stefanel, 2012; Warren, 1983) argue that energy should not be taught as a quasi-material substance, but rather as an abstract concept and the property of a system. Others, on the contrary, (e.g., Duit, 1987) defend the quasi-material metaphor, on the ground that the material-like interpretation leads to the same physical measures. The metaphor makes learning easier at first, but its limitations must eventually be made clear to learners, which may challenge the acquired conceptual models. In biology and organic chemistry, for instance, a consensus in the entire fields accepts that energy is stored in chemical bonds in organic molecules, whereas a chemical bond actually is a form of energy shortage rather than storage (Millar, 2005; Quinn, 2014). This consensus simplifies teaching by providing a convenient conceptual model but it also leaves learners with an imperfect understanding they are not aware of and which can impede further education. Clarification of this metaphor can be conducted with further metaphors, such as a comparison between chemical bonds and Velcro to show the need for energy input in order to break chemical bonds (Quinn, 2014), or direct experience of a bonding force with a molecular modeling kit equipped with magnets such as Snatoms (Muller, 2015).

In the case of energy, the distinction between the two approaches actually pertains to ontological category. Energy as a substance fits the ontological category of "matter", whereas energy as a property of a system can be considered a "process", because it arises from interactions in that system (Chi, Slotta, & de Leeuw, 1994). Learning energy as a quasi-material substance before re-defining it as a mathematical principle requires an ontological shift: reassigning the concept from the "matter" ontological category to the "process" ontological category. Ontological shift is difficult for learners. Learning is easier when conceptions are ontologically compatible (Chi et al., 1994). Accordingly, learning about energy under its scientific definition should ideally be done without using the metaphor of a quasi-material substance in order to avoid categorizing energy as "matter". Furthermore, this ontological discrepancy constitutes a substantial obstacle in the simultaneous learning of two different definitions of energy, societal and scientific. Scientific rigor requires energy to be taught as an abstract principle, whereas society favors an easier and more material definition so that a broad portion of the population understands the concept.

Ontological shift is however not impossible and can be achieved with instruction specifically designed to target conceptual change at the categorical level (Chi, 2008). Furthermore, ontological shift is a normal part of learning, concerning for instance the concepts of sound and heat (Lautrey & Mazens, 2004). Ontological shift is described in these cases as a gradual process of belief revision rather than a sudden transfer (Mazens & Lautrey, 2003).

Beyond their nature, the two definitions of energy also diverge on their properties. A first difference lies in the requirement of a frame of reference. In its scientific definition, energy is an emergent property of the state of objects in a system, and depends on how the system is defined, i.e., what frame of reference is chosen. For instance, one does not think of a lake's water as having gravitational potential energy when one bathes in the lake. However, in a larger frame of reference that includes the lake dam and the river below, one can suddenly think of water as having gravitational potential energy that can be converted into electrical energy by the dam's turbines. Similarly, any inert object on Earth, like a pebble, is in fact packed with massive kinetic and gravitational potential energy in reference to the center of the Milky Way. This energy is never considered, and, for all matters and purposes, does not exist, because this reference frame is never relevant when talking about pebbles. In contrast, under the societal definition of energy, the frame of reference is not defined. When a battery is discharging, the energy it contains leaves the battery and could be measured somewhere else, maybe in thermal or kinetic energy in the device it powers. However, for the owner of the battery, the energy is simply gone, consumed. At larger scales, the same logic applies. When fuel is consumed, the energy it contained leaves the frame of reference of interest to society and is thus considered consumed as well. This leads to another, and major, discrepancy between the two definitions: the conservation of energy.

Under the scientific definition of energy, the law of conservation states that energy cannot be created nor destroyed; it is conserved in all chemical and physical reactions. Only a few scientific principles are established as laws of nature, and the conservation of energy is one them. Without conservation, the scientific concept of energy loses all purpose and can be discarded. Under the societal definition, however, energy is produced, created, or harvested, then used, consumed, or wasted, or even burnt, for instance when "energy" in talked about in "calories". In any case, it does not follow the law of conservation. Even more, the most common mentions of energy under the societal definition are about *saving* energy. Energy, in this view, is elusive, evanescent, and impermanent: quite the opposite of a concept primarily defined by its conservation. This discrepancy stems from society's practical perspective on energy and the absence of a clear frame of reference. For society, energy that is present in the environment but cannot be effectively put to use by an engineering device is not considered energy.

Because the law of conservation of energy is central to the scientific definition of energy itself (see definitions above; Feynman et al., 1965; Poincaré, 1902), the acquisition of this law is seen as the ultimate goal of scientific energy education (M. Bächtold & Munier, 2014; Lee & Liu, 2010; Neumann, Viering, Boone, & Fischer, 2013). A first approach in teaching about energy is thus to start with the law of conservation as a solid foundation, before developing the details progressively (Millar, 2005). A second approach is to follow a progression towards the understanding of the law, through various milestones or aspects of energy, and progressively adapting the vocabulary (Colonnese et al., 2012; Lee & Liu, 2010; Neumann et al., 2013). Milestone concepts include sources and forms of energy, transformation, transfer, dissipation, and, finally, conservation (M. Bächtold et al., 2014; Lee & Liu, 2010; Neumann et al., 2013).

At the macroscopic scale, conservation of energy is not observed because friction, resistance, radiation, and other kinds of dispersion always occur and cannot be perfectly measured. With scientific teaching arguably excessively focusing on idealized frictionless situations as opposed to realistic situations (Nordine et al., 2011), students can easily misunderstand when conservation of energy is applicable. For Bächtold and colleagues (2014), the societal definition of energy does not sufficiently emphasize the principle of energy conservation. In fact, and although it is hard to observe, conservation of energy technically remains true at the societal scale. This is good to keep in mind because it does not only mean that energy cannot be destroyed; it also implies that energy cannot be created. Being aware of the energy conservation principle may thus prevent naïve thinking such as believing in the extraction of energy from perpetual motion machines and lead to the understanding that all sources of energy come in limited, although possibly large, supply. Furthermore, perceiving energy as a property of the state of a system rather than as a quasi-material substance implies that all transfers of energy produce a form of waste: energy cannot be simply extracted from fuels like a fluid; it requires changes to be applied to the system. Citizens applying this thinking and seeking sustainability would be able to question at what time scale a certain energy source may become problematic because of its available supply or waste by-products.

Another discrepancy between the two definitions concerns the scientific concept of entropy, which Duit (1984) argues is the key to understanding the societal definition of energy, thus reconciling both scientific and societal worlds. Simply stated, entropy is a measure of the likelihood of a system's state. With time passing, all unlikely states of matter tend to turn into more likely ones. In other words, according to the second law of thermodynamics, entropy always increases in closed systems. An example may clarify the concept: Ice cubes in warm lemonade always melt and cool the lemonade, but cool lemonade never spontaneously warms up while spawning ice cubes. This is because in the many possible random arrangements of heat in the lemonade glass, there are many more possibilities of heat being uniformly distributed, resulting in cool lemonade, than there are ways to pack heat in some places but not others, resulting in warm lemonade with ice. Lemonade with ice cubes is not a likely state of matter, and time inexorably suppresses it. In the lemonade glass, as in any closed system, energy is quantitatively conserved but its distribution inevitably spreads, flattening differences that could lead to movement, action, or change. This phenomenon, characterized by useful energy gradually turning into useless energy, is called energy degradation.

The notion of energy degradation intrinsically explains the illusory disappearance of energy and thus expresses the law of conservation in its own way: energy is not lost but it loses value. Thus, although entropy and energy degradation may not be known nor understood by lay people, it is experienced so much in everyday life that it hinders learning the law of conservation of energy. Everyday experience and observation of physics phenomena indeed plays an important role in understanding scientific concepts (Vosniadou, 2002). For this reason, Doménech and colleagues (2007) recommend teaching energy with a focus on entropy in order to link the societal and scientific definitions of energy. Similarly, Duit (1984) highlights that scholars in the field of energy supply, rather than thinking of "saving energy", conceptualize in terms of "minimizing energy degradation". Duit argues that energy degradation in realistic situations should be introduced before the law of conservation in idealized systems.

Although seemingly less accurate, the societal definition of energy accounts for complex physical phenomena in simple ways. To account for physical phenomena, like for instance the combustion of fuel, physicists make use of two concepts: (a) energy which always remains constant in a closed system, and (b) entropy which always increases in a closed system and degrades energy. In society, this is generally simplified with a single concept of energy that gets consumed over time. It may be more scientifically accurate to describe energy consumption as energy degradation or as a decrease in lowly entropic energy, but highly entropic energy is arguably not energy at all. This can be seen in the example of the lemonade glass. At first, in the glass, there is more thermal (internal) energy in the warm lemonade than in the cold ice cubes. Because this setting is lowly entropic as compared to a glass of cool lemonade, change occurs and thermal energy is distributed. But would anyone describe the new isolated system, a glass of cool lemonade, as having thermal energy in its own frame of reference? For all intents and purposes, because it is homogenous, the lemonade has no thermal energy: it has no capacity to do work or undergo changes. The lemonade's thermal energy, once distributed, is as irrelevant as a pebble's gravitational potential energy in reference to the center of the galaxy. Therefore, from the point of view of society, it simply does not exist.

Because of its explanatory potential, the societal definition of energy should not be considered flawed or false, rather accepted as a different concept from the scientific concept. Emphasizing the duality rather than ignoring it would encourage metaconceptual awareness

(Vosniadou, 2002) and enable the articulation of the two concepts. This articulation is discussed in the following section.

2.2 Articulating the concepts

Several discrepancies set apart the scientific and societal definitions of energy. Under its scientific definition, energy is a mathematical principle that applies within a defined frame of reference and follows the law of conservation of energy and the second law of thermodynamics. It is an abstract tool invented by physicists that requires a complementary and even more complex concept, entropy, to fully explain physical phenomena. Under its societal definition, energy is a quasi-material substance existing regardless of frame of reference and subject to creation and destruction. The societal viewpoint thus combines energy and entropy into one fuzzy but more intuitive concept.

However, such discrepancies between two definitions or two concepts going by the same name obviously lead to confusion. First, understanding the complexity of the concepts and the subtlety of the differences between them requires knowledge that teachers often do not master (M. Bächtold & Munier, 2014). To make matters worse, the two definitions are often taught in the same class and by the same teacher. As a result, teachers end up having a double discourse which can be described as schizophrenic (M. Bächtold, 2014). Finally, teachers do not realize the difficulty for learners to be faced with two different worlds of energy. As a consequence, learners may resort to the wrong definition when faced with a problem to solve, erroneously integrate information from one concept into the other, or even fail to notice the contradictions altogether and ignore that there are two different concepts, making their articulation impossible (M. Bächtold & Munier, 2014). To attempt to solve these problems, teaching strategies can be derived from the study of concept acquisition and conceptual change.

The first concepts that learners acquire, before formal instruction, are called preconceptions (Vosniadou, 2012), also known as naïve framework theories (Vosniadou, 1994) or phenomenological primitives (diSessa, 1993). Vosniadou (2012) argues that preconceptions constitute cohesive explanatory frameworks, much like the scientific conceptions acquired later through education. Conceptions and preconceptions mature through two different processes: enrichment and revision. Enrichment consists of filling the framework with consistent complementary information without confronting present knowledge. Revision, however, is required when incompatible new information or properties are added to the framework or model. Revision can lead to misconceptions: conceptions that produce systematic patterns of error (Vosniadou, 2002). The classical approach to conceptual change focuses on addressing learners' misconceptions by confronting them with correct information, a process named *cognitive conflict* (Posner, Strike, Hewson, & Gertzog, 1982; Vosniadou, 2012). More recent approaches to conceptual change also consider learners' preconceptions and their progression towards cohesive and consistent scientific conceptions and mental models (diSessa, 1993; Smith, diSessa, & Roschelle, 1994; Vosniadou, 2012). Furthermore, the certainty or uncertainty a learners holds about a conception can also be used as a driver of the learning process, leading new uncertainty and epistemic questions to emerge from introduced certainty (Tiberghien, Cross, & Sensevy, 2014).

Misconceptions are particularly likely in the case of energy because energy differs in nature between its two definitions. Misconceptions arise more frequently when *emergent* phenomena, which occur because of interactions in the system, are perceived as *direct* phenomena, which are caused by a single element of the system (Chi, 2005). Under its scientific definition, energy is an emergent phenomenon: The interaction of physical forces leads to the emergence of energy. Under its societal definition, however, energy is a direct phenomenon, existing independently from the system as a quasi-material substance. Learners do indeed tend to perceive energy as a tangible entity and a direct phenomenon even under the scientific definition (Duit, 1987), as shown for instance in a common misconception about heat and temperature: the belief in the existence of "heat particles" (Chi, 2005).

As a consequence, it is of key importance for learners to integrate the articulation between the scientific and societal concepts, or as labeled by Tiberghien and Megalakaki (1995), the theoretical and applied levels. Without conceptual articulation, students integrate facts from different perspectives or different scales into synthetic models breeding misconceptions (Vosniadou, 2012). Knowing when to apply either the scientific definition or the societal definition of energy is a legitimate component of mastering the concept (Quinn, 2014), and as Solomon (1983, p. 50) puts it, “Pupils must never lose the ability to communicate. It would indeed be a poor return for our science lessons if they could no longer comprehend remarks like ‘wool is warm’ or ‘we are using up all our energy’.”

In sum, learners must necessarily acquire and learn to differentiate two different concepts of energy, which are ontologically incompatible and feature conflicting properties. Energy as a substance could serve as both the basis of the societal definition and a scaffold, or temporary metaphor, for the scientific definition. Specifically designed instruction would later allow science students to conduct an ontological shift of their scientific definition of energy while retaining the societal definition unchanged. Another option could consist of emphasizing the relativity of all scientific theories. The largest ontological gap indeed resides in the definition of energy as an emergent process and the definition of energy as a fiction invented to simplify physics calculations: a gap between a physical reality and a human invention. All concepts, models, and theories of science are ultimately abstract ideas, simplifications, and inventions. They describe and predict nature, but they *are not* nature. Because learners believe a concept to be real, they try to fit all their knowledge about it into a single "true" framework, whereas multiple scientific models of nature can be simultaneously correct yet incompatible. A classical example in physics is the wave-particle duality: quantum-scaled objects (e.g., photons) can be described as both waves and particles, and the accuracy of either model depends on the physical phenomenon considered. If these models were to be synthetized together, they would not provide truth but misconceptions. The two models are thus taught in parallel. Similarly, energy could be taught as a concept that allows the quantification of physical phenomena, and to which a different nature and different properties are given depending on context. The next section shows that energy is also given different units in different contexts, which themselves also lead to confusion and provide obstacles to learning.

2.3 Confusions in energy units

The scientific literature does not seem to question the role of energy units in learning and comprehension. However, as human inventions arbitrarily scaled, units could very well be an available lever to help learning about and understanding energy, especially given the confusion that surrounds them. Domestic energy units sometimes need to be explained in non-educational contexts because they are so poorly understood, for instance in political newspaper articles (e.g., Le Postillon, 2016). Errors may even find their way into academic papers in the field of energy management itself, with for instance "kilowatt-hours" mistakenly written "Kilowatts/hour" (Froehlich, 2009, p. 3), which would mean "kilowatts per hour" and have very little practical sense. Froehlich (2009, p. 4) also mentions "the amount of kilowatts consumed the previous day", which is a non-sense. The author probably means "the amount of kilowatt-hours". The neologism "kilowattevers" coined by an energy user (Strengers, 2011) beautifully conveys citizen's confusion and poor understanding of energy units. The causes of this confusion can be found in the origins and definitions of energy units.

The unit of energy in the International System of Units is the joule (J), defined as the energy transferred to an object by moving it one meter against a force of one newton. In the same system, the unit of energy rate is the watt (W), defined as one joule per second. These units were not designed to express daily domestic magnitudes of energy rates or energy. For instance, a machine wash requires about five million joules. This ridiculously high value is obviously unpractical in daily life, and other units were invented to express magnitudes of energy in home energy and various other fields.

In order to have a unit at an adapted scale, relevant comparisons were sometimes used. When the first engines were created and set in competition with horses, the power of these engines was naturally compared to that of horses, leading to the creation of the horsepower unit. This unit exists in two versions today: the mechanical or imperial horsepower, approximately equal to 746 watts, and the metric horsepower, approximately equal to 735.5 watts. This way, early steam engines and those that followed could readily be compared to horses in terms of, for instance, their profitability. Moreover, this arguably provided an intuitive sense of the engine's power for people familiar with horses and the work they could achieve. It allowed people to imagine the engine at work and its power. Another example of a unit of energy based on a direct relevant reference is the electronvolt. This unit, empirically defined as a specific variation of energy in the charge of an electron, is adapted to and used in particle physics. In this field, the SI unit was much too large to be of practical use, as the value of the electronvolt shows: one electronvolt is approximately equal to 1.60×10^{-19} joules. Other direct equivalent units are used in other fields, such as the barrel of oil or ton of oil equivalent in large energy production and consumption, or the TNT equivalent for explosives. In other cases, units were entirely made up and did not have a previous meaning, like the British Thermal Unit (1 BTU = 1055 J) or the Therm (about 105.5 MJ). The existence of these units highlights that appropriately scaled units are important enough to be invented even when they don't spontaneously arise from intuitive comparisons.

In home energy, the most used units are the kilowatt-hour for amount of energy and the watt for energy rate. As stated above, the watt is the SI unit of energy rate, a measure of energy (joule) per time (second). The kilowatt-hour is based on the watt, multiplied by a unit of time (the hour) and divided a scaling factor (kilo, 1000). The kilowatt-hour can thus be expressed by the following:

$$\text{kilowatt-hour} = \frac{\frac{joule(energy)}{second(time)} \times hour(time)}{1000}$$

This notation shows that the time dimensions cancel out in the equation. The kilowatt-hour thus only quantifies energy, with no time component at all. However, the presence of the word "hour" in the name suggests a time component. Intuition and experience with other units like the kilometer per hour or the oscillation per second lead people to erroneously understand a kilowatt-hour as a kilowatt per hour, whereas a kilowatt per hour does not make sense as a unit of energy rate. Similarly, the absence of a word referring to time in "watt" suggests that a watt refer to an amount of energy, and not an energy rate. The time component in the watt is hidden in its definition as a joule per second. Consequently, not only are the two main units used in home energy confusing in themselves, but they also encourage users to understand one as the other: the watt as unit of energy, and the kilowatt-hour as unit of energy rate.

Ideally chosen names for units should feature the dimensions (energy, time) the units respectively refer to, and they should be at an appropriate scale. An appropriate scale leads to more frequently using one- and two-digit numbers. Using fewer digits is a general recommendation found, for instance, in style guides such as the one followed by the present text, and it is also the principle behind scientific notation. Fewer digits may be easier to process because, even when irrelevant, visible digits cannot be ignored and are automatically processed (Dehaene, 1997, p. 75; Korvorst & Damian, 2008). Also, larger numbers are harder to handle due to a compression of their mental representation (Dehaene, 1997, p. 76). This applies to both energy and time dimensions: the hour may thus be preferred over the second.

Consequently, maybe a new energy unit could be invented for the purpose of domestic use. It would be a big endeavor to promote its use among world citizens, but given the importance of energy for the climate and the economy, it may be an investment worth the effort. Ideally this unit would be worth 18 kilojoules (5 watthours). Given this value, an LED lightbulb would use about 1 unit per hour, an incandescent (60 watts) lightbulb would use 12 units per hour, and a clothes dryer, the most voracious home appliance, would use about 800 units per hour. A typical French household would use about 2.6 kilo-units per day (13 kilowatt-hours) with an annual consumption of about 950 kilo-units. With this unit, the joint manipulation of quantity and rate should be more intuitive and would prevent confusions such as these observed with watts and kilowatt-hours. It is unlikely that the present proposal will lead to the wide acceptance of a new energy unit, but the idea is arguably at least worth considering for further research in energy management and education.

2.4 Conclusion

Energy is a confusing concept. It follows two incompatible definitions (scientific and societal) and is quantified, in its domestic use, with poorly scaled and oddly named units (kilowatt-hour and watt). The two definitions of energy differ on their nature or ontological category: in the scientific definition, energy is a process arising from interactions in a system, whereas in the societal definition, it is akin to matter, described as a quasi-material substance. The two definitions of energy also differ in their properties. In the scientific definition, energy only exists when defined in a specific frame of reference or isolated system, follows the law of conservation of energy, and requires the concept of entropy to explain macroscopic scale phenomena. In the societal definition, energy exists regardless of frame of reference, is not subject to conservation but on the contrary can be produced and consumed, and does not require the concept of entropy. To make confusion worse, the two definitions are often taught in the same class and by the same teacher, who also does not master the concept. As a result, learners acquire misconceptions about energy.

In order to reduce this confusion, energy education can follow either one of two paths. First, focus can be placed on the scientific definition of energy and the law of conservation. Learning about energy this way would prevent an ontological shift from "matter" to "process", and would allow citizens to understand that energy is conserved at a societal scale as well, and realize that no energy source is unlimited and that energy extraction always implies a change in the system. Conversely, energy education can teach energy as a quasi-material substance both for the societal definition and as an introduction to the scientific definition. In this case science students will eventually have to conduct an ontological shift to conceive of energy as a process rather than matter. In any case, it must be made clear that another concept of energy exists in society, and the differences between the two concepts must be emphasized. Failing to understand and learn the conceptual duality of energy leads to misconceptions. As a result, learners may be unable to understand the scientific definition of energy because they apply the societal definition to scientific situations. Similarly, scientifically fluent people could lose the ability to communicate about energy in society because they hold on to its scientific definition.

Confusion also arises from energy units, because they are not appropriately scaled for daily use, and also because their names evoke dimensions they do not contain (e.g., kilowatt-hour being independent of time). New units of energy could easily be invented, but spreading their use worldwide would be a real challenge.

This chapter highlights that society seeks and needs an intuitive concept of energy assorted with intuitive units, in order to empower citizen with the ability to make energy-related decisions. The scientific definition of energy cannot be used for this purpose because it is not adapted. To account for macroscopic scale phenomena, the scientific definition requires other concepts such as entropy. Furthermore, learning the scientific definition of energy is difficult, because learners intuitively think of energy as a material-like substance, rather than an emergent process. Shifting from one conceptualization to the other is a difficult task. Seeking to educate all citizens to a scientifically sound definition of energy could prove counter-productive as compared to teaching an appropriate societal definition of energy to all, and a scientific definition to some. In practical applications, a metaphorical material definition of energy should be used for most people, if and only if the metaphor does not lead to erroneous assumptions. In home energy management for instance, energy can be thought of as a tangible resource and quasi-material substance. Furthermore, in the absence of adapted units, quantification and communication of amounts of energy remains a problem, which will be discussed in the next chapters.

3

MONITORING ENERGY

Using less energy is important today because saving energy helps slowing down climate change, because the major source of electricity in the world remains fossil fuels. Of course, keeping fossil fuels consumption in check will only delay the problem of climate change. Technology must adapt. Renewable energy sources will be used, such as wind and solar. These sources, however, suffer from a lot of variability over time. Winter gives less solar energy than summer, cloudy days less than sunny days, and wind is also variable. Overall, these energy sources cannot be controlled as well as coal and nuclear, and production undergoes more variation. Already, electricity is more expensive during peak hours, although energy providers have a lot of control over its production. In the future, the price of electricity is very likely to vary even more according to its availability. It is also possible that temporary energy shortages will happen more often than today. Citizens will have to learn to be adaptable.

In response to this situation, various actions can be undertaken, one of which is to reduce energy consumption. A few people choose to reduce their energy consumption to the maximum, refusing any entertainment or comfort that requires burning fuel or consuming electricity. The masses, however, seem not to want to stop living their life and abandon their comfort in order to save energy. Most people are willing to make an effort to save energy, but this effort is limited. How can citizens be helped to save energy and maximize the impact of their actions? How can they be helped to figure out when and where an effort will be worth it?

3.1 Home Energy Management Systems

The electric grid is the largest machine on Earth. Millions of appliances are connected to it, and each receives energy on demand. Electricity can travel tens or hundreds of kilometers between production site and consumption site. Until the recent dawn of ubiquitous computing, that is, the proliferation of electronic devices, the grid was merely a massive electromechanical machine. Energy use was hard to monitor and thus hard to control. But in the recent years, many electronic systems have been developed to measure and communicate energy use to energy providers as well as inform energy users. The tools destined to users are named Home Energy Management Systems (HEMS). These devices allow energy users to understand, control, and manage their energy use. HEMS are different from smart meters: smart meters collect information, mainly for the energy provider, whereas HEMSs make it available to inhabitants (Van Dam, Bakker, & Van Hal, 2010). HEMSs typically dispense information, comments, advice, and/or rewards. Using such tools and understanding their contents is not innate for people. A lot of research has thus been conducted on their design, and a lot of different tools have been developed. In this chapter will be discussed how energy management tools can be categorized, according to their goals and properties, and what are the theories behind them.

The main of goal of HEMS is to help people reduce their energy use. However, before rushing in discussing the means to achieve this goal, it needs to be further detailed. The grid is (as yet) not equipped with massive batteries to store energy, thus all energy produced, if not used, is wasted. The ultimate goal of energy saving is thus to prevent producing energy at the production site. This is one of the goals of smart meters, like Linky currently in installation in 1many households in France. It allows energy providers to acquire information about energy use, and thus to optimize their production. However, production cannot be easily and efficiently adjusted to fit steep variations in energy needs. Factories have maximal outputs as well as optimal production ranges, and varying their output takes time, work, and energy. Also, many sources of energy do not allow much variation, if at all. Solar and wind energy are for instance not in control of the provider. Thus, energy providers seek to cut peak energy use, in order to produce more constantly and avoid energy waste, a strategy known as peak shaving. For this purpose, they offer energy plans with cheaper electricity off-peak. HEMS can be used for that purpose as well, by informing users of peak hour times or of the current price of electricity. This is the purpose for which the Energy Orb, a glowing display, was used. Originally designed to monitor stock markets, it was repurposed for energy monitoring by an energy provider who distributed the Orb to its major customers: large industries. The Orb, combined with attractive off-peak energy prices, drastically reduced peak energy demand (Holmes, 2009).

Another goal of energy management comes from the evolution that the grid is undergoing. The classic model consists of few large centralized production sites such as coal or nuclear plants distributing over long distances, but renewable sources reshape this landscape. Scattered small-scale production sites will be more and more numerous, bringing energy down to a local rather than national scale. Citizens are for instance able to produce a lot of their own energy with solar panels. Also, renewable energy sources are typically unable to respond to demand variation, being dependent on sunlight, wind, or tides. When only renewables are used, production control escapes the hands of the provider, so it is users who must adapt their consumption. The roles are reversed. Batteries may provide a buffer to counteract this dependency, but batteries are expensive. Consequently, user adaptability would reduce the need for batteries, saving work, money, and resources, and making ambitious projects possible. It is in fact such a project, a habitat with limited resources, which inspired this thesis. Monitoring energy consumption as well as production would allow lowering energy use when consumption is high and/or when production is low.

Finally, reducing energy use obviously also consists of reducing total energy use, lowering the baseline, in order to lower the overall production needs. This is what is most commonly thought of when thinking of energy saving, but it is important to note that it is not the only important goal, as explained above.

In order to achieve these goals, HEMSs have been designed according to a variety of strategies and perspectives. For Yun and colleagues, (2013), nine forms of interventions can be undertaken with HEMSs, namely: education, advice, self-monitoring, comparison, control, reward, goal setting, engagement, and communication. Similarly, Pierce, Odom, and Blevis (2008) describe seven available strategies: offering behavioral cues and indicators, creating social incentive to conserve, connecting to material impacts of consumption, encouraging playful engagement and exploration with energy, projecting and cultivating sustainable lifestyles and values, facilitating discussion and raising public awareness, and stimulating critical reflection. A single tool can aim to achieve multiple goals, or do so unintentionally. Karlin (2011) highlights that tracking and learning can be achieved with the same HEMS, are both important, and lead to different effects. Focusing on either one of these goals should influence designers' choice between presenting the data either in real-time and aggregated or as a history with much detail. The numerous strategies and intervention techniques constrain and define the practical properties of HEMSs, which provide even more complexity when these technical variables are taken into consideration. Van Dam and colleagues (2010) list seven variables that define each HEMS. These seven variables show the variability that can be found in HEMSs as objects and the difficulty to compare them to one another. First, they include the purpose of the HEMS, which can be total energy saving or peak shaving as discussed above. Also, different forms of energy can be monitored, for instance electricity or gas, and can also include water. Feedback can also be provided at various levels, from the level of a single appliance, as do simple wattmeters, to the entire household. Visualizations at larger scales, such as this of a neighborhood or a city, do also exist, more as a form of art (e.g., *nuage vert*, Holmes, 2009). Van Dam and colleagues (2010) also consider physical variables of HEMSs, such as their location in the house, the option for controlling energy use from the device itself as opposed to simply monitoring it, and the modality of physical interaction, which can cover the whole range of human-computer interaction tools. Finally, the type of feedback is considered, which varies widely with the goals and approach of the device, both in terms of their contents and of their form.

Given the complexity of HEMSs and the long list of possible ways to categorize them, another type of categorization needs to be proposed. Here, I divide HEMSs in two categories: displays requiring active engagement, and displays requiring only passive attention. Those requiring active engagement all have in common that they tend to receive much attention when they first arrive in a house, and are progressively abandoned. The mere presence of a device indeed makes people really aware and engaged (Pierce, Fan, Lomas, Marcu, & Paulos, 2010), but this initial interest vanishes with time. HEMSs drift into the background and users stop looking at them and paying attention to them. They are not effective anymore in four months of time, after which users revert to previous behaviors (Van Dam et al., 2010). Only a specific niche of users, the most motivated towards green behavior, can get the best results. Thus, HEMSs requiring active engagement can only be effective in the long term if they are thought of as educational devices, which must durably change user's behavior before being inevitably left out. HEMSs of the other type, which do not require active attention, are called ambient displays or ambient visualizations.

3.2 Ambient displays

Ambient displays are based on the idea that HEMSs should not be designed like control panels in airplane cockpits. Pilots dedicate their full attention to their instruments, which is of primary importance because the life of their passengers depends on it. Home energy management is much less critical, and is not a task that is conducted with users' full attention. Energy users are not professionals sitting all day in front of a display and a range of buttons to control their home appliances. In fact, HEMSs should be used as little as possible and stay in the background, so as not to interfere with users' life and pursuit of happiness. Given this perspective, ambient displays typically provide visual information that does not need ocular focus: they glow. I have already mentioned one successful ambient display, the Glowing Orb, which originated from Massachusetts Institute of Technology's Tangible Media Lab and enabled huge energy savings by informing factory employees of the cost of energy by glowing red during peak use. Holmes (2009) reports that previous attempts by the energy provider to inform their customers, via calls or emails, had not been successful because these means of communication required dedicated time and attention. On the contrary, the Orb provided information than could be gathered in a glance.

Other similar glowing HEMSs have been created, such as Greeny (Wever, van Kuijk, & Boks, 2008) and Wattson. Both feature, in addition to the glowing ambient display, a digital display providing numerical information like instantaneous or cumulated energy use. Ham and Midden (2010) tested a similar ambient display in the form of a color glow, varying from green to red according to instantaneous energy use, in comparison to a numerical display in watts, in a task of temperature management of various rooms of a house. The ambient display led to lower energy use, particularly when participants conducted a simultaneous numerical distraction task, which suggests that it can effectively be perceived in the background of another task. All these devices are quite similar, using a soft glow to communicate information in the background of people's attention. As such, they could be used in the long term and constitute an energy management tool used daily.

Displays featuring a notch of art have also been created, such as a radiator made out of lamps, which intrinsically glows proportionally to the heat it emits (Gyllenswärd, Gustafsson, & Bang, 2006) or the Power-Aware Cord, which is an electric extension cord that glows blue, varying in intensity with the current flowing through it (Backlund et al., 2007).

Despite the argument made here that ambient displays could avoid the fate of other HEMSs and remain long effectively in use, a sample of users provided with Power-Aware Cords for two months showed a progressive lost in interest in the device, and eventually likened it to a Christmas light (Löfström & Palm, 2008). This questions ambient displays' long term usefulness. Another limitation of ambient displays is that they provide very simple information with little precision. It is very hard to determine exactly how much energy is being used at an instant simply from the color of a glowing light, but completing ambient displays with numerical displays can compensate for this limitation.

3.3 HEMSs as educational devices

As described above, HEMSs which do not feature ambient displays only manage to capture users' attention for a limited period before being ignored. Their initial use can actually be explained by a short-lived curiosity which momentarily compensates for the unpracticality of regularly having to check on such a device. Rather than being thought of as professional tools, HEMSs must be thought of as educational devices which will be used a lot during an initial learning period and be used occasionally afterwards. De Vries (2001) provided a topology of computer-supported learning which partly applies to HEMSs. This framework comprises eight pedagogical functions, which are built upon four main theoretical viewpoints. The eight functions are (1) presenting information, (2) dispensing practice drills, (3) actually teaching, (4) catching learner's attention and motivation, (5) providing an exploration space, (6) providing an environment for the discovery of natural laws, (7) providing an environment for the discovery of abstract domains, and (8) providing a communication space between learners. The four theories are behaviorism, cognitivism, constructivism, and situated cognition. Some HEMSs could really be analyzed within this framework, being actual games. Gamberini and colleagues (2012) for instance developed a game linked to household's energy consumption which provides tailored advice, has been well accepted, and yielded nice results. Shiraishi and colleagues (2009) also used various persuasion techniques in their game/system. They rewarded positive behavior with instant positive feedback, drawing on behaviorism, and used strategies of social comparison as well. However, most HEMSs are actually not thought of, by their designers, as educational software. For this reason, the topology of computer-supported learning does not exactly fit the categorization of HEMSs. It is proposed here to categorize them according to the three theoretical approaches that correspond to HEMSs as tools providing feedback in order to induce long-lasting behavior change. The three approaches are behavioral, social, and cognitive.

3.3.1 Behavioral approach

Persuasive technology, which includes HEMSs, is defined as aiming to induce a voluntary change in individuals' attitudes or behaviors (IJsselsteijn, de Kort, Midden, Eggen, & van den Hoven, 2006). Thus if the goal of HEMSs is to change users' behavior, behavior change technique should be most appropriate. This technique is largely based on Skinnerian learning, also known as operant conditioning. DiSalvo, Sengers, and Brynjarsdóttir (2010), for instance, mention a "Skinnerian" way of changing behavior in persuasive technology. Simply stated, operant conditioning consists of providing either a reinforcer or a punisher following a given behavior from a person. Reinforcers increase the likelihood of the behavior occurring again in the same context, and punishers have the opposite effect. In its simplest form, operant conditioning does not require comprehension; reinforcers or punishers are provided, and behavior changes. This mechanism seems simple enough, but leads to very effective behavioral change programs when considering precisely the details of behavioral learning. Three important variables must be considered: the specific behavior targeted, the choice of the stimuli used as reinforcers, and the consistency with which the technique is applied. Proper behavior change can only be conducted with tight control over these three variables, and this control is typically impossible in the case of HEMSs.

First, increasing the likelihood of a behavior occurring again with positive reinforcement requires a reinforcer to be provided right after the behavior was conducted. This must not happen several minutes after the fact, but within a few seconds. In some cases, electric bills are considered reinforcers, and they are provided long after the behavior is conducted. Real-time data is more efficient (Smeaton & Doherty, 2013), but users still need to see their device in order to receive this reinforcer, which may occur long after the behavior is conducted, thus any reinforcer provided is disconnected from the corresponding energy-saving behavior. Moreover, even if a HEMS could provide a reinforcer at the moment a behavior is conducted, for instance by sending a notification on the user's phone, the HEMS could only reinforce simplistic approximations of energy-saving behaviors, such as turning off appliances, because only the consequences of behaviors can be monitored, not the behavior themselves. If HEMSs could determine what good energy saving behavior are, then energy management could simply be done by a computer. At its best, a system based on the behaviorist approach could only mimic the old incentive to "turn off the lights" so often repeated to children: a rigid and simplistic rule, radically different from the smart management required given the complexity of the energy mix that fuels the grid, with no real change of habits except for a few trivial behaviors (e.g., Pierce et al., 2010). Finally, the use of punishers to avoid wasteful behavior is not accepted by users, as exemplified with the rejection by users of feedback consisting of frowning faces (Holmes, 2009).

Second, reinforcers need to be tailored to each individual; there is no such thing as universal reinforcers. Tailoring reinforcers is possible when a psychologist or educator carries out a well-planned behavioral change technique, but impossible to do with HEMSs. In fact, reinforcers must be unavailable in the learner's environment, and it may be very hard for an electronic device to provide such a reinforcer. A solution might be gamification. Games can provide reinforcers that are only available within the game, such as achievements, equipment, bonus levels, game money, or even just high scores. However, the use of games depends on users' will to literally play them, and, for this, games need to be and remain fun. Building a fun game increasing energy-saving behavior and reaching a wide range of users would be a real challenge, and before it is achieved it can be considered utopian.

Third, reinforcement must first be applied with consistency, and then given a specific schedule. If this is not accurately done, behavior change will not last in the long term. HEMSs do not implement such long term changes, because applying an appropriate reinforcement schedule is difficult and requires a properly designed and personalized behavioral change intervention. In the absence of such schedules, achieved energy-saving behaviors disappear after the HEMS ceases to be used.

The behavioral approach thus suffers from a lot of shortcomings. It may seem simple and efficient, but in reality important behaviors cannot be targeted, learning varies across people, and changes do not last.

3.3.2 Social approach

The behavioral approach presented so far basically consists of building habits. However, social psychology describes that people engage in behaviors not only out of habits, but also because of what they believe they should do or believe is normal to do. In other words, behavior can be derived from a rational choice or the adoption of a norm (Froehlich, Findlater, & Landay, 2010).

Rational choice, in this view, includes individuals' attitudes (their predisposed state of mind regarding a value), which influence behavior according to the theory of planned behavior (Ajzen, 1985). Behavior could thus be changed by changing attitudes. Cornelissen, Pandelaere, and Warlop (2006) tried to induce ecological behavior change after attitude change. They noted that attitudes towards ecological behavior, which are more important than attitudes towards ecology itself, are negative. They managed to induce a change in attitudes but found no results on behavior.

The adoption of norms follows another mechanism, one of social comparison. People tend to conform to social norms and this can be used to motivate them to use less energy (Smeaton & Doherty, 2013). Because data on energy use are actually hard to come by, energy users want comparisons both with themselves and with similar households. In both cases, comparison provides context and sets the norm (Froehlich et al., 2012). Pierce and colleagues (2010), for instance, found that users base their goals on their baseline energy use, which they try to maintain, but not to lower: the baseline becomes the norm. Other people's energy use can also become the norm because of social comparison, which can be implemented very easily. For instance, Ayres, Raseman, and Shih, (2013) show that providing households with some neighbors' energy use along their bills yields small yet long-lasting energy savings. Similarly, Midden and Ham (2009) found that adding social comparison to factual energy data led to higher energy savings. They moreover advocate for the use of a social agent, that is, a virtual character, as a means to provide the data, so as to foster social effects such as comparison. Social comparison thus constitutes a useful tool to foster energy saving, but it also raises problems such as competition between people, and privacy issues such as accountability and blame which user see as a possible consequence of monitoring and displaying people's home behavior (Froehlich et al., 2012).

Another viewpoint involving social norms can be found in Strengers (2008). She postulates that the limit to further reducing energy use lies in the social norms of household behavior. Normative behaviors, that is, what is considered normal, change a lot over time and geographical location. Cleanliness, for instance, is not defined by natural human standards but is a social construct and varies across places due to historical reasons. The culture of neatness and cleanliness, in the Netherlands, is an example of a local idiosyncrasy that has its roots in the need for immaculate conditions for the manufacture of butter, and which survived centuries (Pye, 2014). Strengers (2008) argues that energy users stay within the boundaries of normative behaviors when they try to save energy, which limits the scope of possible savings, while energy feedback systems do not try to challenge the norms. Strengers (2008) suggests two ways to accomplish this. First, feedback systems should lead to more discussion, which allows norms to be locally redefined. For this reason, HEMSs should be centrally located in the house and noticeable. Second, HEMSs should provide alternative norms, for example by suggesting unconventional choices to a given situation.

The social approach to energy-saving behaviors is promising, with long-lasting effects due to social comparison and possibly even longer lasting changes arising from the re-definition of normative behaviors. However, social feedback must nonetheless be understood by users, graphs and charts must be appropriately compared, and successfully using the social approach thus requires representations suited for human cognition.

3.3.3 Cognitive approach

A third approach consists of easing information processing in order to increase understanding and learning. This approach is based on cognitive research. DiSalvo and colleagues (2010) refer to this as passive persuasion, data provided without comments, as opposed to strong persuasion, data accompanied by reinforcers or punishers such as smiley or frowny faces. The latter type of persuasion corresponds to the behavior change techniques described above. Empowering users with the ability to understand their energy use data, as opposed to simply have them change their behavior, provides them with much finer control and could lead to higher and more precise savings (for instance during peak use). This knowledge could also extend to other domains where energy is used, such as transportation or professional settings, and would also grant citizens the freedom to make informed choices rather than obey imposed social norms and follow naively trained habits. Also, computers cannot be programed to satisfactorily control home appliances. Currently, appliances can be designed to be remotely controllable, but intelligent users are still needed. Human cognition is needed. For this reason, behavior change techniques can only be limited to a certain amount of savings, after which it would reach its limits.

Designing HEMSs that would foster understanding and mastery of energy use is a hard task. What is needed is to present information, or feedback, in a comprehensive way and in accordance with cognitive theories. Feedback can be provided in a huge variety of ways. First must be determined *what* is displayed: should it be financial cost, energy used, or environmental effect? As discussed in the first chapter of this thesis, the choice of a unit is in itself a difficult question. Fischer (C., 2008) describes that feedback varies also in terms of frequency and duration, breakdown or agglomeration of data, comparisons with norms and previous data, as well as presentation format, and argues that regarding the comprehensibility and appeal of text or graphics, "the devil is very often in the details", leading most HEMS projects to be unable not tackle these problems. Similarly, after listing ten design dimensions for energy use feedback, Froehlich (2009) concludes that "the ways in which to most effective build interfaces around these data to reduce consumption is an open research question and one that involves psychology and human-computer interaction."

Most HEMSs indeed do not take into consideration a user's ability to easily process the information displayed. They are described as not efficient, not ergonomic (Wever et al., 2008). The ability to read and understand graphics, graphicacy, is an under-explored topic (Bétrancourt, Ainsworth, de Vries, Boucheix, & Lowe, 2012), and, accordingly, only a few studies in the field of energy use feedback concerned themselves with the question of graphical representation of feedback. For instance, Wilhite, Hoivik, and Olsen (1999) compared different types of graphs presented along paper bills, namely a social comparison graph based on a line graph versus a bell curve. Energy users responded positively to both feedback, but the differences in energy saving between the two groups were small and significativity was not assessed. Similarly, Egan (1999) presented different forms of graphs to energy users. Between different forms of lines, a bell curve, and a distribution graph with little houses, energy users preferred the distribution graph, and understood but did not like the bell curve. Current research on representations in the field of HEMS does not go beyond the scope of such studies.

Regarding graphical representations in HEMSs, much is however to be questioned. Their main problem is that they use arbitrary and varying scales. Centimeters, square centimeters, pixels, and colors do not have a consistent meaning. Graphs typically scale up or down according to the data they contain. For instance, in order to accommodate a peak on a graph of recent energy use, the whole *y* axis representing energy use would be compressed so that the peak would not end up literally "off the chart". As a result, the rest of the data would be flattened, leading to an impression of more constancy, and the meaning of slopes would change. Because of such variations, users are unable to compare graphs from day to day solely from their appearance. They always need to check the scale of the axes and mentally compute the differences instead of relying on the visual impression the graph produces. A similar example of this phenomenon can be experienced in Google Maps, which provides the elevation profile for bicycle routes. Because the scale changes with each route's length and difference in height, cyclists cannot compare routes between one another or associate a physical slope to a slope on the profile. Thus, the profile only indicates when the road ascends or descends, but not if the ascent will be too steep for a cyclist. Theories of grounded cognition discussed in the next chapter, stipulate that cognition is based on mental simulations of perceptions and actions (Barsalou, 2008; M. Wilson, 2002). Given this perspective, graphs are objects that can be manipulated as images, mentally layered on top of one another for comparison, but not scaled out of initial proportions. The mind does not process them by extracting the data they contain like mathematics software; as a matter of fact graphs communicate data better than tables because of their intrinsic analog (visual-spatial) properties, which only exist in Euclidian space. Destroying this space with inconsistent transformations, or software magic tricks, prevents the analog processing of graphical representations, leaving the graph to be simply processed like a table.

Providing energy use data according to the human mind's processing abilities can lead users to learn and understand the data, which could lead to long-term savings if this is combined with behavioral techniques to build habits and social techniques to change attitudes and provide motivation. The effect of HEMSs being based on data feedback, the presentation of data in a cognitively ergonomic fashion is central to their success.

3.4 Guidelines and conclusion

HEMSs could lead to energy savings, but they are really hard to consider in their individual properties and to compare to one another. In the long term, attention to displays vanishes, and with it vanish good habits and savings. Also, different goals are to be considered, for instance general savings and peak shaving, making difficult the comparison of different HEMSs. Nonetheless, reviews find that using HEMSs results in 5 to 15% of energy saving, with an average around 7% (Darby, 2006; Faruqui, Sergici, & Sharif, 2010; C. Fischer, 2008).

The behavioral, social, and cognitive approaches together provide a set of rules that HEMS design should follow. Feedback should be given in real-time, include color-coding, and be built in comparison to a baseline and contextual information as well as self-comparison. The data should be available on mobile devices, anywhere, any time, to provide feedback as close to behaviors as possible, and easy access to historical data should be included for self-comparison (Smeaton & Doherty, 2013). Accurate and individual data should be accessible effortlessly by users, energy-saving behavior should be made to look normal to provide normative influence, and the links between causes and effects should be underlined. Mobility and gamification should be encouraged (Zapico, Turpeinen, & Brandt, 2009). Finally, HEMSs must take all inhabitants into consideration, and invite to changing behaviors that are otherwise seen as non-negotiable (Strengers, 2011).

HEMSs use various theories and methods that are all valid and potentially useful. Games can be created, automatic feedback can be provided, social comparison can be put to use. However, in all cases external representations of energy use are going to be cognitively manipulated by users. This cognitive manipulation must be ergonomic, easy on the brain. In all cases the representation of magnitude must be taken into account. The next chapter discusses the mental representation of magnitude in estimation and comparison from the viewpoint of numerical cognition (e.g., Dehaene, 1992) and grounded cognition (Barsalou, 2008; M. Wilson, 2002).

4

MAGNITUDE SENSE AND GROUNDED COGNITION

Mathematics, arithmetic, and simply numbers are abstract ideas that humans process with a brain designed for perception and action. Knowledge about the mechanisms that enable numerical cognition is of central importance in the improvement of teaching about mathematics and numbers, as well as in the facilitation of energy management. In energy management, nearly all tasks consist of mental comparisons of amounts of energy. Whether confronted with an unusual energy bill, the purchase of a new appliance, or a political environmental decision, a citizen needs to consider what the amounts means in reference to other amounts or to a mental reference: Did we use a lot of energy this month? Is this refrigerator efficient or voracious? Can I run all these appliances at the same time? Are the savings planned for each household realistic? In each case, a mental comparison of amounts of energy with a relevant reference solves the issue. The dimension concerned by such mental comparisons is magnitude. Magnitude is the notion of quantity or amount of something that makes it quantifiable and subjectable to quantitative comparisons. Size, mass, and brightness for instance all have magnitude. Magnitude exists on a continuum and can be an attribute of concepts that cannot be measured or counted precisely. For instance, brightness can be easily visually compared but is nearly impossible to quantify with human eyes. The present chapter describes how the use of perceptual and motor processes explains human cognitive abilities via mental simulations as described by theories of grounded cognition (Barsalou, 2008; M. Wilson, 2002). Emphasis will be placed on numerical cognition and magnitude processing (Dehaene, 1992, 1997; Walsh, 2003).

4.1 Perceptual and motor grounding of cognition

Grounded cognition is an umbrella terms that describes a variety of viewpoints on human cognition, all of which share a common denominator: the rejection of the idea, born at the dawn of electronic computing in the twentieth century, that the human mind is information-processing software computing abstract symbols. Grounded cognition rejects this idea and argues that the body shapes and affects the mind. The term "embodied cognition" describes roughly the same theories as "grounded cognition", but, as highlighted by Barsalou (2008), "embodied" suggests that bodily states are required for cognition or are always of central importance in cognition. Describing cognition as grounded rather than embodied is a way to distantiate the general perspective of grounded cognition theories from a subset of more extreme viewpoints only focused on bodily states. Thus, the two terms describe very similar perspectives. Some authors distinguish "grounded", "embodied", and "situated" as different components of congruent theories (M. H. Fischer, 2012; Pezzulo et al., 2013) but this distinction is not followed by all authors. Finally, the terms "grounded" or "embodied" cognition should not be taken to mean that a form of "non-grounded" or "disembodied" cognition simultaneously exists in the mind. On the contrary, grounded cognition theories defend that there is no cognition but grounded cognition.

According to grounded cognition theories, the basis of human cognitive abilities lies in the use of sensorimotor processes for general cognitive functions. In this view, sensorimotor processes are mobilized for mental simulations, enabling cognition to handle absent objects, revive the past, anticipate the future, imagine other locations or objects never encountered, and give rise to abstract cognitive abilities such as language processing and problem solving (Barsalou, 2008; Dijkstra & Post, 2015; M. Wilson, 2002). Performing mental simulations consists of mentally carrying out tasks or imagining actions and situations without actual or complete sensorial input or motor output. Mental simulations can be conscious and voluntary, like when one imagines oneself in a fictive or past situation, but more importantly, can be unconscious covert mental simulations constantly generated to sustain cognition.

As indicated by neuroimaging evidence, mental simulations make use of the same cortical areas as full-blown sensorial and motor processes (for reviews, see Barsalou, 2008; Jeannerod, 2001; Martin, 2007). This activation remains contained within the brain and little to no signs of it can be readily observed in the body. This explains why mental simulations could only be fully scientifically studied after brain imagery technology enabled their observation as neural activations. Some views of grounded cognition have specifically focused on *perceptual* simulations, such as Barsalou's (1999) *perceptual symbol systems* as well as earlier accounts of visual imagery (Kosslyn & Pomerantz, 1977), whereas other views centered around *motor* simulations (Jeannerod, 2001).

Mental simulations parsimoniously explain how primate brains acquired abstract thinking without evolving a new organ, by making use of existing structures and mechanisms. This constitutes what biologists call an exaptation: the development of a novel function by an already existing feature. Accordingly, no specific brain area exists to handle abstract concepts. On the contrary, abstract concepts are represented by neural networks in the sensory and motor cortex, as indicated by neuroimaging evidence (Martin & Chao, 2001). Neuroimaging studies further indicate that training, learning, imitation, evaluation of the consequences of one's actions, and consciousness of agency are the consequences of covert mental simulations in motor areas of the brain (Jeannerod, 2001). A review of other neuroimaging studies can be found in Barsalou (2008).

Mental perceptual and motor simulations of action further explain the strong link between body and mind revealed by the branch of research focused on embodied cognition (see for instance A. D. Wilson & Golonka, 2013; M. Wilson, 2002). This research shows that seemingly irrelevant body states affect cognition. For instance, Elder and Krishna (2012) showed that when simply presented with a visual advertisement for a graspable product, people spontaneously simulate grasping the product. If the orientation of the product does not match their dominant hand, or if their hand is busy holding something, then intention of purchase decreases. This shows that although grasping is irrelevant for abstract processing of intention of purchase of the product, the actual covert cognitive mechanism determining intention of purchase makes use of a mental simulation of grasping.

Finally, another property of cognition with mental simulations is that cognitive activities can be offloaded into the environment (Kirsh, 2010; Risko & Gilbert, 2016), and cognition can thus be referred to as *distributed* between the mind and the environment, across individuals, and across space and time (Zhang & Patel, 2006). This basically constitutes the opposite of generating mental simulations of real phenomena with perceptual and motor processes. The transaction from mind to environment and vice-versa is made possible by the use of common processes: perceptual and motor. Cognitive offloading into the environment was for instance identified in a study using the video game Tetris, in which participants, when time-pressured, preferred rotating the pieces on-screen, rather than mentally, in order to find matches between pieces and available slots (Kirsh & Maglio, 1994). In this case, because mental rotation has a finite angular speed (Shepard & Metzler, 1971) whereas Tetris turns pieces instantly, offloading may have been the faster and certainly least cognitively effortful option.

In summary, the model of a cognition based on motor and perceptual mental simulations is both powerful regarding the explanation of phenomena and parsimonious with respect to the brain structures it requires. The model of cognition using mental simulations also suits earlier findings in cognitive science and classical models of cognition, such as the model of working memory featuring a phonological loop and a visuospatial sketchpad (Baddeley & Hitch, 1974). In the examples given here, auditory and verbal mental simulations can be understood as the phonological loop and visual simulations as the visuospatial sketchpad (M. Wilson, 2001). However, theorizing the ability to use mental simulations with existing perceptual and motor processes is more parsimonious than theorizing an entire separate cognitive architecture.

4.2 Theory of embodied mathematics

The process by which motor and perceptual processes lead to abstraction can also be understood as a succession of experientially grounded metaphors nested within one another. Lakoff and Johnson (1980) argue that a single ability, this of using metaphors, allows action-oriented and physically grounded beings to scaffold their cognition into higher levels of abstraction. Given this viewpoint, humans understand abstract concepts and live by metaphors rooted in bodily experience. Lakoff and Johnson (1980) oppose the objectivist view that meaning is disembodied in favor of the experiential view according to which meaning is specific to a person and arises from the available cognitive tools they have and according to their experience of life.

Metaphors in Lakoff and Johnson's (1980) sense may be seen to be mental simulations, motor and perceptual, applied to other contexts than their original context. For instance, many orientational metaphors pervade human life and discourse, such as "more is up" and "less is down". These metaphors are not only observed in language (e.g., "a raise in income") but also provide an experiential ground, here perceptual, for thinking about abstract quantities such as money or energy. Also, these metaphors have a physical, experiential basis, namely here that the more is added to a pile of objects, the higher the pile becomes, for example "energy is a pile of wood". Thinking about numbers with this "pile" metaphor possibly leads to a vertically oriented visual-spatial mental representation of numbers, as was experimentally observed (e.g., Shaki & Fischer, 2012). Similarly, the metaphor "energy is a fluid" provides an experiential ground for understanding the concept of energy, leading to metaphors in language (e.g., "energy flows from the battery into the circuit") but also to associated heuristics (e.g., expecting that the battery will become empty), thus enabling both language and abstract thinking.

Mathematics is often regarded as the most noble and specifically human skill, the pinnacle of abstract thought, but complex mathematics have nonetheless also been explained by a series of metaphors or mental simulations. The major work on this topic was undertaken by Lakoff and Núñez (2000) who present the *theory of embodied mathematics*. The authors follow the metaphorical understanding of human cognition of Lakoff and Johnson (1980) mentioned above, and apply it to the domain of mathematics. Numerical magnitude is understood, via metaphor, as a collection of objects, a length on a measuring stick, or a motion along a path, among others. Various linking metaphors also connect the different branches of mathematics together, and metaphors of bodily action such as movement are used across mathematics to describe mathematical objects such as functions (Núñez, 2006). The theory of embodied mathematics even outlines a philosophy of mathematics and tackles metaphysical questions such as the ontology of mathematical objects, which are described as metaphorical human constructs grounded in bodily experience (Lakoff & Núñez, 2000).

4.3 Mental simulations in numerical cognition

Mental simulations, as described in theories of grounded cognition (Barsalou, 2008; M. Wilson, 2002), parsimoniously explain many processes of numerical cognition, the acquisition of numerical skills, and the nature of mathematics. Relying on perceptual and motor processes, mental simulations do not require dedicated or abstract mechanisms in the brain. As Hubbard and colleagues (2008) argue, regarding numerical cognition: "uniquely human abilities arise from abilities that have been conserved across phylogeny". Mental simulations enable abstract cognitive processing via progressive reduction of both motor action and perceptual input during physical processing, as well as offloading of cognition into the environment. Typical examples are counting, multiplying, and stacking additions. Counting can be explained by motor mental simulations of counting on one's fingers (Domahs, Moeller, Huber, Willmes, & Nuerk, 2010). Initially, a learner counts on her fingers aloud and with full hand movements, then she progressively reduces her movements until they become hardly noticeable while lowering her voice to a whisper, then she reduces them further until all overt movement and voicing disappear and only mental simulations remain

(M. Wilson, 2002). Such mental simulation of finger counting may remain the basis of counting in adults. A functional imaging study showed activation of hand motor circuits in a task where items were to be serially ordered (although this did not only concern number items but also letters; Andres, Seron, & Olivier, 2007). Also, finger counting habits seem to influence numerical cognition in adults by contributing to grounding numbers in a spatial arrangement (Wood & Fischer, 2008). Another basic mechanism making use of mental simulations in numerical cognition is multiplication. Everywhere in the world, multiplication is acquired verbally by rote learning via recitation and repetition (Dehaene, 2001). Once multiplication facts are acquired as short phrases, one can mentally multiply, e.g. four times four, simply by mentally simulating saying the phrase "four fours are sixteen" and only say out loud "sixteen". Similarly, once a learner masters stacked addition on paper, she can progressively learn to conduct the same process in mind, with visual mental simulations, as long as the number symbols to be manipulated remain manageable. Combined with verbal mental simulations of rote learned addition phrases, these mental simulations enable mental addition of multi-digit numbers.

4.4 Mental simulations in magnitude processing

Like other processes of numerical cognition described above, magnitude processing can be explained by perceptual mental simulations. Magnitude processing is acquired through the body in infancy, extends from continuous quantities to discrete quantities and then to numbers (Bueti & Walsh, 2009). Mental simulations of perceptual experience can explain this evolution.

4.4.1 Magnitude comparison of continuous dimensions

Processing magnitude is a deeply rooted ability which allows comparisons of continuous dimensions. This ability has been observed in infants and across the animal kingdom on continuous dimensions such as size, brightness, and loudness (Dehaene, 1997; Leibovich, Katzin, Harel, & Henik, 2017). In human adults, comparison, i.e. the discrimination of two stimuli varying in the magnitude of a common property, has been studied in psychophysics. Psychophysics research showed that the difference perceived between two stimuli varies logarithmically with the actual difference (Henmon, 1906). In other words, the more similar two stimuli are, the more difficult it is to discriminate them. This property known as the Weber-Fechner law is an important feature of perceptual comparisons and which demonstrates the involvement of perceptual processes whenever it is observed.

Mental comparisons of continuous dimensions also involve perceptual processes and analog representations when they are conducted from memory. Moyer (1973) accordingly describes that an "internal psychophysics judgement" takes place when the size of two animals, pictured from memory, is mentally compared. The process of generating analog mental images in order to carry out cognitive tasks with perceptual processes such as comparison from memory was first named "imagery". Imagery was defended as an explanatory construct in cognition (Kosslyn & Pomerantz, 1977) which could enable various tasks such as scanning images for attributes (Kosslyn, 1973; Kosslyn, Ball, & Reiser, 1978) and mental rotation of objects. Shepard and Metzler (1971) showed that in a task of mental rotation of shapes, response time was proportional to the angle of rotation, indicating that rotations were performed analogically rather than computationally or propositionally. Although the name "imagery" may suggest to some that only visual imagery existed, motor imagery was also described (Jeannerod, 1995) and other sensorial modalities were considered (e.g., Kosslyn & Pomerantz, 1977).

Perceptual rather than categorical processes were shown to underlie mental comparisons of magnitude. When imagery was proposed, a debate in cognitive psychology opposed analogical to propositional views of human cognition. For imagery supporters, imagery could carry out cognitive processes analogically, and theorizing propositionally organized processes for the same tasks was unnecessary and unparsimonious. For supporters of a propositional cognitive system, the mind was organized as a network of propositions (Fodor, 1975; Pylyshyn, 1973, 1984). Observing and comprehending a scene consisted of establishing true and false propositions about the scene using abstract concepts as fundamental bricks of thought. "The elephant is bigger than the mouse" used the concepts of "elephant", "mouse", and "big", as well as an abstract relationship of comparison. This perspective is similar to the way linguists understand and analyze language, and it also corresponds to the way computers are programmed to "think": abstract objects with abstract attributes. The comparison of the size of objects was shown to be by default conducted analogically, with mental images, but could also be conducted with categorical processes when the compared objects had been overlearned as belonging to different size categories (Kosslyn, Murphy, Bemesderfer, & Feinstein, 1977). This is also applicable for instance to the comparison of an elephant, a typical "large" animal, to a mouse, a typical "small" animal. The category alone would lead to an inference about which animal is larger. In this case however, the task is arguably not mental comparison of magnitude but rather categorical identification and was possible because objects were compared one by one and only two categories existed. Therefore, whether conducted from stimuli or from memory, comparisons of continuous magnitude involve perceptual processes and continuous dimensions, i.e., mental simulations. The role of categorical mental process is nonetheless to be considered when categories are more relevant than magnitudes.

4.4.2 Magnitude comparison of discrete number of objects

Mental comparisons can also be conducted on number of objects, in which case it also involves perceptual processes and continuous mental representations (Dehaene, 1997). Number of objects, for pre-verbal minds like infants and animals, cannot simply be summarized in a number made out of symbols like words or Arabic numerals. Without such a tool, discriminating between two similar numbers of objects requires a different cognitive mechanism. Research converges towards the idea that discrete non-verbal discrete amounts larger than four (Feigenson, Dehaene, & Spelke, 2004) are mentally represented with continuous mental representations identical to mental representations of continuous dimensions (Gallistel & Gelman, 2000). These representations intrinsically feature imprecision, also named scalar variability or noise, due to their continuous rather than analog nature. The generation of a continuous mental representation from discrete stimuli is conducted with an accumulator (Gallistel & Gelman, 2000; Meck & Church, 1983). Dehaene (1997) also defines animals' counting mechanism as an accumulator, metaphorically described as a primitive water tank used by some neurologically impaired Robinson Crusoe (Dehaene, 1997, p. 28). The island survivor fills the tank with an identical amount of water for every individual cannibal he sees, in order to visualize their total number as amount of water in the tank. Water, a continuous quantity, is used in the metaphor rather than stones or sticks, discrete objects, because without the ability to manipulate discrete symbols like numerals or units of any kind, animals have to resort to summing and estimating magnitudes on a continuum, represented by water in the metaphor. This mechanism can't show the difference between 99 and 100 predators, but it is sufficiently accurate to discriminate between "none", "few", and "many". This suggests that minds without language do not possess a "number sense" but rather a "magnitude sense" which is put to use in number estimations.

The existence of a "number sense" in opposition to a "magnitude sense" is currently the topic of a scientific debate in the community of numerical cognition (Leibovich et al., 2017). The keystone of the debate is the impossibility to display a discrete number of objects without intrinsically displaying associated continuous magnitudes such as length or surface area. For example, six red dots on a white background induce more red surface area than five dots do. A line of six dominoes is longer than a line of five dominoes. It is difficult to manipulate numbers independently form magnitude. Some evidence nonetheless shows that infants are more sensitive to change in number of objects (discrete) than change in magnitude of a continuous dimensions, suggesting the existence of a number sense independent from a magnitude sense early in infancy (Libertus, Starr, & Brannon, 2014). Conversely, the Weber-Fechner law is observed in comparisons of number of objects, suggesting the processing of a continuous, rather than discrete, representation. The only case where the Weber-Fechner law is not observed is in subitizing, that is, when small discrete quantities (1, 2, 3, and up to 4) must be identified (Dehaene, 1997). Subitizing is a different process from counting and may involve the recognition of geometrical configurations (Mandler & Shebo, 1982) or be a process of its own (Dehaene, 1997). Uncertainty remains around the nature of subitizing and the existence of an innate number sense or magnitude sense. Evidence nonetheless shows the use of continuous perceptual mental representations in magnitude comparisons.

4.4.3 Magnitude comparison of numbers

Magnitude comparison of numbers, e.g. Arabic numerals, is also conducted with a continuous (or analog) visual-spatial mental representation. This mental representation often takes the form of a number line (Dehaene, 1992). The three mental representations underlying numerical abilities are described in Dehaene's (1992) *triple-code model.* The sense of magnitude comes from the analog magnitude representation, in which magnitude is represented on a number line. This representation, being continuous rather than discrete, includes imprecisions, as portrayed in Dehaene's (1997) metaphorical description of the accumulator. Evidence shows that the number line obeys the Weber-Fechner law of psychophysics. First, determining the larger of two numbers is achieved faster when the numbers are "further apart" on the number line, i.e., when the numerical difference between them is larger. This constitutes the *distance effect* (Moyer & Landauer, 1967). Second, the number line is compressed at large magnitudes, i.e., larger numbers are more difficult to discriminate (Dehaene, 2003).

The two other mental representations in the triple-code model allow other processes than mental comparisons. One mental representation is the auditory mental representation which allows for instance rote learning of the sequence of natural numbers, i.e. counting, and rote learning of multiplication tables and other basic arithmetic. The other mental representation is symbolic, specifically in Arabic numerals. This mental representation allows multi-numeral operations such as stacked additions, subtractions, multiplications etc., as well as all of arithmetic and basic algebra. The symbolic representation in Arabic numerals is not used in mental comparisons but merely allows access to the corresponding analog magnitude representation, which is the one used in mental comparisons. Mental comparisons of number are not conducted as subtractions with the symbolic-Arabic representation, as a computer would do: "subtract B from A, if result is positive then A is larger than B".

4.4.4 Magnitude comparisons as mental simulations

In conclusion, the argument is made here that magnitude processing is conducted from the basis of perceptual mental simulations, as evidenced by the pervasiveness of perceptual processes in magnitude processing and given the framework of grounded cognition presented above. In this view, mental representations of magnitude of continuous dimensions, discrete amounts, and number all consist of perceptual mental simulations of analog continua, except in the case of small discrete amounts, in which case another perceptual process elicits subitizing. The following section presents more details on the perceptual mechanism involved in the mental comparisons of these perceptual mental simulations that constitute mental representations of magnitude.

4.5 The common spatial mechanism of magnitude comparison

As detailed above, all magnitudes are mentally represented by perceptual processes as analog continua. However, this does not describe the properties of the comparison mechanism. The first property of magnitude comparisons is that they are conducted by a single system in the brain. A Theory of Magnitude (ATOM; Walsh, 2003) describes that all magnitudes are processed in a single generalized system, placing space, time, and quantity processing together in the same brain location. This idea is very well supported for instance by Lourenco and Longo (2010) who show that infants who are shown large objects of a certain color associate the color not only with large objects but also with objects of great number and objects lasting longer in time, as if "large" concerned all dimensions—space, time, and number. The common magnitude processing system leads to another property that was uncovered using the *size congruity paradigm*, based like the classic Stroop task (Stroop, 1935) on mental interferences: typically, two numerals of different physical size are presented, and the task is to quickly identify the physically larger numeral, ignoring the numerical value. When physical size and numerical value are not consistent, interference occurs as the two magnitudes are processed in the same system in the brain, leading to longer response time

(Henik & Tzelgov, 1982). Cohen Kadosh and Walsh (2009) review findings regarding the existence of a centralized magnitude processing system in the brain, concluding that evidence remains in support of the ATOM theory. It seems Dehaene's (1997) metaphorical Robinson Crusoe did not bother digging out different accumulators to count different things, and uses the same for all purposes.

A second property of magnitude comparisons is that they seem to intrinsically require spatial mental representations, not just any analog perceptual mental representations. This is best demonstrated by the existence of Spatial Numerical Associations (SNAs) (M. H. Fischer & Brugger, 2011). SNAs were first identified in parity judgements as the Spatial Numerical Association of Response Code (SNARC) effect, characterized by faster responses to stimuli of small magnitude with the left hand and faster responses to stimuli of larger magnitude with the right hand (Dehaene, Bossini, & Giraux, 1993). Abundantly replicated (Wood, Willmes, Nuerk, & Fischer, 2008), the effect was also observed with musical stimuli (Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006), and with full-body motion responses given on a dance mat (U. Fischer et al., 2016). SNAs are due to the spatial orientation of the number line with smaller numbers on the left and larger numbers on the right, and shows that "number line" is not a mere linguistic metaphor but corresponds to a spatially oriented mental representation. Similarly, Lakoff and Núñez (2000) describe the spatial metaphor as a core feature of mental numerical representation rather than a mere figure of speech. SNAs can also be observed in vertical responses, possibly because of the spatial metaphor "more is up" (Shaki & Fischer, 2012). SNAs were also shown to be linked to finger counting (M. H. Fischer, 2008), highlighting that magnitude representation is grounded in space via the body (M. H. Fischer & Brugger, 2011). Many other findings from neuropsychological, neuroimaging, and behavioral research highlight the fundamental connection between spatial processes and magnitude representations (de Hevia, Girelli, & Macchi Cassia, 2012; de Hevia, Vallar, & Girelli, 2008; Thompson, Nuerk, Moeller, & Cohen Kadosh, 2013). Consequently, regardless of the modality by which a magnitude is first perceived, for instance with auditory or haptic senses (Gimbert, Gentaz, Camos, & Mazens, 2016), a spatial mental representation seems to always be generated in order to process and compare magnitude.

Many mental representations of magnitude can exist, as long as they are spatial. The number line, so central to theories of numerical processing e.g., (Dehaene, 1992), is an idealization of the complex reality. Two arguments support this claim. First, the number line is a product of formal education in arithmetic and numerical symbol manipulation. Pupils use a variety of forms of number lines (Siegler & Opfer, 2003) which evolve from logarithmic number lines to linear number lines (Siegler & Booth, 2004), gradually expanding in both directions to include larger numbers and negative numbers, and increasing in precision to include fractions (Siegler & Lortie-Forgues, 2014). In the evolution of the representations, early intuitive forms are progressively inhibited, while new formal forms are constructed (Laski & Dulaney, 2015). Second, magnitude can be represented on other spatial representations than a number line. For instance, when the mental representation of a clock face was prompted in an experiment, an inverted SNARC effect was observed because of the placement of numbers on the clock face (D. Bächtold, Baumüller, & Brugger, 1998). People can thus resort to other representations that fit a certain situation and skip the number line. In this view, mental comparisons of number magnitude are in fact perceptual mental comparisons of spatial mental simulations of objects, either segments on a number line or other continuous spatial objects.

In conclusion, the perceptual nature of magnitude comparison, the strong connection between space and magnitude, and the existence of a single centralized magnitude processing system in the brain together suggest that when a mental comparison is to be made from any two stimuli or memory items, a spatial representation of both magnitudes is constructed, possibly on a number line, on which perceptual processes conduct the comparison. Comparisons of amounts of energy thus require spatial mental representations of these amounts.

4.6 Conclusions

Given the framework of grounded cognition, all of human cognition can be understood as resting on perceptual and motor mental simulations. Mental simulations also enable mathematical and numerical cognition, as well as magnitude processing. Magnitudes are processed by mentally simulating spatial continua, using perceptual processes. The comparison of mentally simulated spatial continua is conducted using a perceptual mechanism akin to the "internal psychophysics judgement" described by Moyer (1973). Because magnitude processing intrinsically requires spatial mental representations, and that mental representations are perceptual mental simulations, it follows that the external representations used in teaching and communication of magnitude should be spatial as well. They would lead to spatial mental representation which can be readily used in magnitude comparisons. Research should thus determine whether learning magnitudes with spatial external representations is possible, and whether it facilitates magnitude comparison. The practical implications of the role of mental simulations in cognition could also be explored, in order to unveil new ways of supporting learning and magnitude processing. Applying such advances to the domain of energy education and management could help citizens sensing the magnitude of energy, and enable them to make easier and better decisions at the various scales in which energy is used in society.

5

EFFECTS OF EXTERNAL REPRESENTATION ON LEARNING AND COMPARING MAGNITUDES OF ENERGY[1]

5.1 Introduction

In the context of climate change and of the energy crisis, citizens will be held more and more responsible for their energy consumption. In order to keep their energy consumption low and in tune with the fluctuating energy production of renewable sources, citizens may resort to various strategies, such as identifying both efficient and wasteful appliances, keeping consumption below certain set goals or in the range of their neighbors', or adapting consumption to fluctuating local production (Froehlich et al., 2010). These energy management tasks and strategies involve learning and memorizing the energy consumption of a variety of appliances, performing mental operations such as estimating sums and differences, and mentally comparing energy consumption of (sets of) appliances. Such energy management tasks will be as ubiquitous as electric appliances and executed in what may well be described as a distributed cognitive environment (Zhang, 1997; Zhang & Norman, 1994) simultaneously involving internal mental and external representations of energy consumption. Research should not only determine how energy users can be educated today, but also how

[1] This chapter has been written as a manuscript, (Galilée, Schwartz, & de Vries, in preparation), to be submitted for publication in a journal. The study was conducted at the California State University of Chico, USA. Best appreciation to Neil H. Schwartz for making this project possible.

this topic may be addressed in school environmental courses, as curricula will adapt to train energy-conscious citizens.

Energy consumption is a numerical magnitude: a quantity that can be represented as a number and subjected to "greater than" comparisons. Mass, length, pitch, duration, price, and number of objects are other examples of numerical magnitudes. In educational material, feedback displays, or any other medium, numerical magnitudes can be presented with a variety of external representations. Although a number of terms exist in different fields of study—textual versus pictorial (Mayer, 1997), linguistic versus graphical (Stenning & Oberlander, 1995), sentential versus diagrammatical (Larkin & Simon, 1987), symbolic versus iconic (Peirce, 1906), and descriptive versus depictive (Schnotz & Bannert, 2003)—the present argument will focus on the distinction between *symbolic* and *analog* external representations. Symbolic external representations are made out of symbols such as Arabic numerals, and represent by virtue of convention. They bear no resemblance to what is represented (Peirce, 1906). In modern societies, the quasi-universal symbolic external representation of magnitudes is numerals, in the Arabic or another system. By contrast, analog external representations, such as bar charts, represent by similitude, visual or otherwise, to what is represented. Because many numerical magnitudes encountered in daily life, such as length and size, are more or less directly perceivable by the senses and in particular visually, an intuitive way of representing such magnitudes is graphically through matching spatial relations. For instance, the length of an object can be represented by the length of a line or a bar. As a consequence, analog external representations are cases of intrinsic representation (Palmer, 1978), i.e. the represented and representing world share a same inherent property, in this case transitivity (if $A < B$ and $B < C$, then $A < C$). Additionally, energy, as a magnitude, is largely intangible: invisible, silent, and weightless. Thus, choosing an external representation from perceptual similitude is impossible, and a deliberate choice of external representation is required. Energy management systems use for instance bars and pie charts, horizontal lines, bell curves (C. Fischer, 2008), color hue (Wever et al., 2008), and glow intensity (Gustafsson & Gyllenswärd, 2005; Gyllenswärd et al., 2006).

The current study aims at establishing whether there is an effect of external representation on learning and mentally comparing numerical magnitudes of energy consumption.

5.2 Theoretical background

There is abundant research on mental comparisons in the *presence* of external representations of the magnitudes to be compared. However, much less is known about comparing magnitudes in the *absence* of their external representations, i.e., mentally comparing magnitudes only after the retrieval of a representation from memory. Learning and comparing numerical magnitudes in the absence of an external representation supposes mental representations for both processing and storage. A mental representation activated or generated for the purpose of mental processing is called a *transient* mental representation, and a mental representation stored in memory for an extended period of time is called a *persistent* mental representation.

5.2.1 Magnitude comparisons

Magnitude comparisons were first studied in psychophysics in the 19^{th} century (Fechner, 1860). Researchers described human ability to compare dimensions such as color hue, pitch, and line lengths (Henmon, 1906). The main finding from this research is that perception follows Weber's law: perceived difference between stimuli relates on a logarithmic scale to the actual difference. Increased actual difference was found to ease comparisons and lower response times (Henmon, 1906). Thus, in perception, the more similar two entities are on a certain dimension, the longer it takes to decide which entity is larger on that dimension. The same observation was made in number comparisons: Comparing two numbers is faster when their numerical distance is larger (*i.e.*, comparing 2 to 9 is faster than comparing 5 to 6; Moyer & Landauer, 1967). Moyer & Bayer (1976) coined the term *distance effect* for the relation between distance and speed of comparison. The presence of such a distance effect in comparing both dimensions of stimuli and numbers is best explained by common mechanisms: the perceptual processes described in psychophysics are also used, at least partially, in number comparisons, and the representations used must be analog, sharing properties of visual stimuli, rather than discrete like abstract symbols. Similarly, Shepard & Metzler's (1971) mental rotation experiments revealed a correlation between angle of mental rotation and response times, and interpreted it as evidence of the use of analog mental representations. Using analog mental representations in number processing implies a conversion of numerals into physically comparable analog representations of magnitudes (Moyer, 1973). An extensive field of research on the processing of numbers opened to explore these mechanisms and led to the development of cognitive models shedding light on the transient representations of magnitude processing.

5.2.2 Transient representations of magnitude

According to the currently most accepted model of numerical processing, Dehaene's *triple-code model* (1992), numbers are processed in three possible transient mental representations: (1) a visual Arabic number representation, (2) an auditory verbal word representation, and (3) an analog magnitude representation akin to a number line. A number line is a spatial representation used in mathematics and consisting of a graduated straight line with every point representing a real number. According to the triple-code model, the use of a particular internal representation depends on task demands. The visual Arabic number representation is used to manipulate Arabic numerals. The auditory verbal word representation is used to say and hear number words as well as recall arithmetic facts learned by verbal rote. And finally, the analog magnitude representation is used in estimations and comparisons.

In theory, different transient mental representations specifically fit different tasks. In practice, however, all three mental representations can be used jointly. For example, a single complex mental calculation can first involve mentally structuring the calculation with the visual Arabic number representation, then accessing addition and multiplication tables with the auditory verbal word representation, and finally comparing the end result to a reasonable estimate with the analog magnitude representation (see Trbovich & LeFevre, 2003, for the role of phonological and visual working memory in calculations).

The analog representation of number is particularly important, because it enables the actual processing of semantic meaning, *i.e.* magnitude. The use of an analog mental representation was found to be spontaneous, automatic, and independent of external representation (Dehaene, 1992) or task relevance (Dehaene et al., 1993; Hinrichs & Novick, 1982). However, this representation can take various forms, both across and within individuals. The analog mental representation is Iranians, writing Eastern Arabic numerals from right to left, seem to use an inverted number line (Dehaene et al., 1993). This suggests that different analog mental representations of magnitude are used across cultures, just like different visual numerical representations are across regions (e.g., Arabic vs. Kanji numerals), and different verbal word representations across languages (e.g., /twɛlv/ vs. /duz/). Moreover, an individual may have several different analog representations of magnitude and use them depending on context or task. For example, analog magnitude representations such as a clock-like number circle (D. Bächtold et al., 1998) or a vertical line (Rusconi et al., 2006) have been induced instead of the number line through specific task demands. Developmental research also shows that children's analog mental representation of magnitude evolves with age, from a logarithmic number line to a linear number line (Siegler & Booth, 2004), gradually expanding in both directions to include larger numbers and negative numbers, and increasing in precision to include fractions (Siegler & Lortie-Forgues, 2014). In the evolution of the representations, early intuitive forms are progressively inhibited, while new formal forms are constructed (Laski & Dulaney, 2015). The number line is thus an acquired mental tool progressively crafted during childhood tolerating other coexisting analog mental representations.

Furthermore, analog mental representations of magnitude seem to always rely on spatial continua. This is true in algebraic tasks, where the use of visuospatial working memory is a key to the proper understanding of number and is associated with higher algebra performance (Hurst & Cordes, 2017). The use of spatial analog mental representations also applies to auditory dimensions such as pitch height (Rusconi et al., 2006) and in pre-verbal minds such as infants and animals, who do not know numbers but process magnitude spatially (Brannon, 2006; de Hevia et al., 2012). In fact, the theorization of the number line as a spatial analog mental representation was in great part due to the discovery of the Spatial-Numerical Association of Response Codes (SNARC) in parity judgments (Dehaene et al., 1993), which highlights the strong connection between space and magnitude. The SNARC effect is characterized by faster responses to relatively large numbers with the right hand and to relatively small numbers with the left hand, suggesting a spatial and lateralized mental representation of number. Abundantly replicated (Wood et al., 2008), the effect was also found with musical stimuli (SMARC; (Rusconi et al., 2006), as predicted by Walsh's (2003) theory of magnitude (ATOM). In ATOM, a SQUARC effect is expected on any quantity. Beyond hand laterality, the effect was also found with full-body movements and responses given on a dance mat (U. Fischer et al., 2016). Many more findings from neuropsychological, neuroimaging, and behavioral research highlight the fundamental connection between spatial processes and magnitude representations (de Hevia et al., 2008). This might be explained by the fact that they share the property of transitivity. Spatial analog mental representations can thus take many forms based on spatial dimensions, among which the number line is merely the most commonly observed. In summary, Dehaene's (1992) analog transient mental representation is in fact, more precisely, a *spatial* transient mental representation.

5.2.3 Persistent representations of magnitude

Memorial comparisons involve mentally comparing entities in their absence (Moyer, 1973) and thus require retrieving persistent mental representations from memory. The internal psychophysics model (Moyer, 1973) suggests that memorial mental comparisons of dimensions are achieved via imagery, for example comparing the size of two animals with the eye of the mind (Jamieson & Petrusic, 1975). In Moyer and Bayer (1976), participants learned nonsense syllables in association with circles of different sizes. Prompted with two nonsense syllables, they then conducted memorial mental comparisons of the size of the corresponding circles. Observing a distance effect in response times, Moyer and Bayer concluded that both transient representations used in memorial comparisons and persistent representations stored in memory are spatial representations, hence the 'distance' effect. Kerst and Howard (1978) went further in the characterization of persistent representations. From their research on area and distance estimations, they proposed a re-perceptual model of memory for continuous dimensions in which stimuli are first processed as sensorial experience, then encoded as such in memory, and accessed in memory as re-perceptions, partly via sensorial processes. In this view, persistent representations are holistic images of the original percepts. These models, entirely based on imagery, raised criticism in an era dominated by computational (Pylyshyn, 1973, 1984) and propositional (Fodor, 1975) models of human cognition, and whose proponents viewed imagery as a marginal cognitive process. In response, Paivio (1975) adapted the original paradigm (Moyer, 1973) in various experiments and concluded that persistent representations for memorial comparisons are indeed analog, whereas propositional representations are used in other processes such as naming entities. This view is consistent with Paivio's (1971) dual-coding theory postulating both imagery and verbal processes. Propositional and analog processes were further investigated by (Kosslyn et al., 1977). They tested composite propositional-analog models in a series of experiments involving memorial mental comparisons of stickmen drawings which differed in size (analog dimension) and belonged to either one of two size categories (propositional dimension). Participants learned both actual size and category label before conducting mental comparisons. Kosslyn and colleagues (1977) proposed a race model to explain the findings of both a size and a category effect. The race model postulates that propositional and analog processes run in parallel, and the faster process, depending on the particular comparison at hand, resolves the task. Barsalou (1999) interpreted Kosslyn's theory of imagery (1980) as defending the existence of amodal persistent representations in long-term memory, with perceptual images only temporarily existing as transient representations. However, Kosslyn (1980, pp. 106–111) argued that persistent representations may consist of non-holistic images, thus analog representations, possibly associated with propositional representations. Thus, models of mental comparisons involving imagery and internal psychophysics have been challenged and completed but analog persistent representations remained at their core.

The more recent *grounded cognition* perspective (Barsalou, 2008) completes and extends imagery theories, in particular with the widely supported mental simulations hypothesis (M. Wilson, 2002). Mental simulations are mental reactivations of formerly perceived stimuli, processed as if they were the actual stimuli (M. Wilson, 2002). Barsalou (1999) proposes for instance that memory is made out of perceptual symbols. In a nutshell, perceptual symbols are stored components of sensory and motor activations that can be dynamically reactivated to reenact past experiences as well as generate simulations of new experiences. The mental simulation approach does not limit its application range to visual or perceptual imagery and simulation of action; it also explains higher cognitive functions such as symbol manipulation. These cognitive functions are considered to be acquired skills (Barsalou, 2003) much like propositional processes are in the race model (Kosslyn et al., 1977). Thus, in this view, both analog and propositional (but not abstract) transient representations can be generated at will, and may become persistent representations. Recent research continues to support this claim. Denis (2008) found that mentally comparing the distances between landmarks on an island that had only been verbally described led to a distance effect in the response time patterns. Thus, these memorial comparisons were conducted using the mental image of the island, an analog representation, rather than the original verbal description. This shows that mentally generated transient analog representations can indeed become persistent analog representations.

Accordingly, spatial analog external representations seem more efficient and appropriate than verbal external representations for storing spatial visual information such as the size of objects, animals, geometrical shapes and stickmen, or their relative location. As mentioned above, Dehaene's (1992) triple-code model states that spatial representations are also spontaneously generated in the case of numbers. The question remains open in the case of invisible physical properties such as energy consumption, which possess no proper visual spatial external representation. Hence, the particular choice of external representations may influence learning, retrieval, and processing of corresponding numerical magnitudes.

5.3 The current study

In the present study was tested the hypothesis that a magnitude of energy consumption is stored in memory in a persistent representation that shares the properties of the original external representation. An alternative model would be that magnitudes are always stored in a spatial magnitude representation regardless of initial external representation. In order to address this question, an experiment involving magnitude processing was designed. First was a phase of learning and recalling, which prepared for the second phase involving comparisons from memory.

The first phase involved learning then recalling the energy consumption of several appliances with either a symbolic or a spatial external representation. The symbolic external representation consisted of Arabic numerals; the spatial external representation consisted of a bar and provided no scale. Therefore, two of three codes from Dehaene's model (1992) were solicited: the visual Arabic number representation in the symbolic condition, and the spatial analog magnitude representation in the spatial condition. However, in both cases, spontaneous translation from one code to the other was feasible. For instance, in the symbolic condition, the Arabic numerals could trigger a magnitude and its corresponding spatial mental representation. In the spatial condition, a number between 0 and 100 could be attributed to the bar, the latter value corresponding to the top of the screen. Consequently, the persistent representation stored in memory could not be determined without a following experimental phase revealing more data. Recall was expected to be slightly higher in the symbolic condition because numerals are discrete and allow precise recall, whereas graphical bars are intrinsically continuous.

The second phase consisted of memorial comparisons of energy consumption. The energy consumption of a single appliance was compared to a single other one (one-by-one comparisons). Such comparisons involve retrieving energy consumption from memory, obtaining a transient spatial representation, and conducting the comparison. Regarding external representation, comparisons were expected to take longer in the symbolic condition as compared to the spatial condition because it involves a symbolic-spatial translation step at the time of retrieval. Comparisons varied in distance between the two appliances: they were close, medium, or far apart in terms of magnitude of energy consumption. The two types of comparisons were dissimilar in nature and therefore designed to be analyzed separately. Since comparisons take place using a spatial mental representation (Dehaene, 1992), a distance effect was expected in both the symbolic and the spatial condition on both response time and accuracy. Regarding a possible interaction, the distance effect might be slightly stronger in the symbolic condition, because appliances which are far apart in energy consumption (easily decipherable difference) require less precise spatial mental representations to be generated from the numerals. Close distance appliances on the contrary require more precise spatial mental representations which, in the symbolic condition, might take longer to generate, and could be less accurate.

In order to test a more ecological and complex situation, a second type of comparison task was designed which involved two sets of three appliances (three-by-three comparisons)., Given the experimental materials and the outcomes of pilot studies, it was conjectured that three-by-three comparisons would present the appropriate difficulty as compared to two-by-two or four-by-four comparisons for instance. These comparisons involved retrieval of energy consumption from memory, estimation of the total energy consumption on each side, and finally comparison. According to Dehaene's (1992) triple-code model, both estimation and comparison are conducted using the spatial mental magnitude representation. In consequence, three-by-three comparisons rely on the same mental representations as one-by-one comparisons but involve extra steps. Therefore, three-by-three comparisons were expected to take longer than one-by-one comparisons, but not necessarily to be less accurate. The same effects of external representation and distance as for one-by-one comparisons were expected for three-by-three comparisons, with comparisons taking longer in the symbolic condition as compared to the spatial condition, and taking longer for close comparisons as compared to far comparisons. In addition, it was expected to find that comparisons at small distances lead to even larger response times when learning is conducted with a symbolic external representation as compared to a spatial external representation. The distance effect on response time was expected to be stronger in the symbolic condition as compared to the spatial condition because close comparisons may require several iterations of the generation of a spatial mental representation in order to achieve the required level of precision.

Mental manipulation of both numbers and spatial mental representations could be involved in the experimental tasks in both conditions. Accordingly, numerical and spatial skills were measured in order to control for their effect.

5.4 Methods

Two independent variables were manipulated in the present study. First, external representation of appliances' energy consumption was either symbolic (with numerals) or spatial (with bars), creating two experimental conditions. Experimental conditions differed only in the materials used in learning and recalling energy consumption. Second, distance refers to the numerical difference between energy consumption of appliances when comparing them. Comparisons were arranged in three distance groups: close, medium, and far, and distance was varied within participants.

5.4.1 Participants

A hundred and four undergraduate students were recruited via a sign-up sheet in a Californian State University and rewarded with extra credit. Six participants failed to complete the tasks in time and their data were discarded. Of the 98 remaining participants, 78 percent were female, 22 percent were male. Age ranged from 18 to 52 years, *M* = 22.2, *SD* = 4.09.

5.4.2 Materials and apparatus

Four tasks were designed for the experiment: (1) a task of familiarization with the appliances, (2) a learning task, where participants learned the energy consumption of the appliances, (3) a recall task about these energy consumption values, and (4) a task of mental comparison of energy consumption.

For the purpose of the experiment, eight electric appliances with a silhouette, a name, and a description were invented and presented in the familiarization task, followed by a quick questionnaire about them. The silhouettes were carefully drawn to be equally different from one another, so that no silhouette was more recognizable than the others. All silhouettes were in black and white and contained the same number of black pixels (within 1 percent). The names and descriptions were chosen not to evoke particularly low or high energy consumption (Figure 1; see also appendix 9.1). Materials in the familiarization task were the same across conditions.

Name: Sock-and-Roll
Used in: Bedroom
The Sock-and-Roll folds, organizes, and dispenses the inventor's socks according to the forecasted weather.

Figure 1. An example of an appliance.

The learning task involved repeated presentation of eight computer-simulated flashcards. The cards were presented on the screen one at a time, with an appliance's silhouette on the front side and the corresponding energy consumption on the back, and the cards flipped when clicked. Here, experimental conditions varied in the external representation of energy consumption. According to experimental condition, energy consumption was expressed either as a number, in the symbolic condition, ranging from 147 to 701 ($M = 365$, $SD = 181$) or as a vertical bar of corresponding height, in the spatial condition, ranging from 24 mm to 113 mm ($M = 5.90$, $SD = 2.93$). Figure 2 reproduces an example in both conditions.

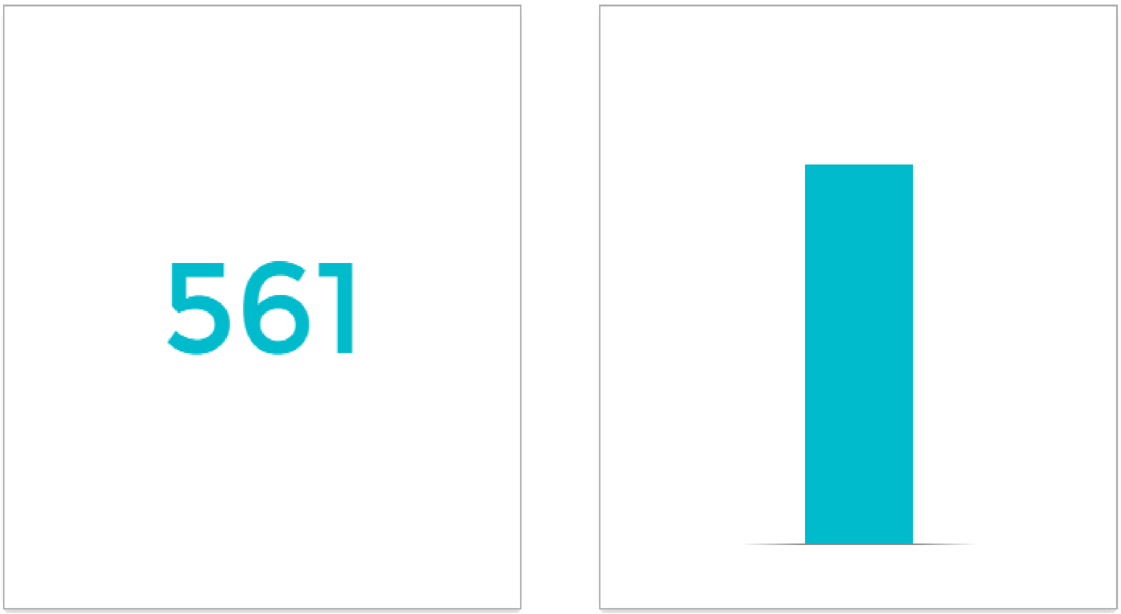


Figure 2. Symbolic (left) and spatial (right) external representations of an appliance's energy consumption.

The numerical values and corresponding bar lengths were chosen after pilot studies hinted that magnitude could be ignored by adopting a mere ordinal processing strategy. In consequence, the perception of a regular step in energy consumption between the consecutive appliances was prevented by varying the seven numerical distances. A logarithmic function was used, which technically created a mathematical sequence, but one that is unlikely to be detected by the untrained eye. The values chosen were increments of 25% of the preceding value (logarithm function to base 1.25, see Table 1).

Table 1. Energy consumption in digits and millimeters

Appliance	1	2	3	4	5	6	7	8
Value in digits	147	184	230	287	359	449	561	701
Size in millimeters	24	30	37	46	58	72	90	113

The presentation order of the flashcards followed a confidence-based repetition design, or CBR (Cohen, n.d.), i.e. depending on the learner's knowledge confidence on a declarative scale from one to five. In CBR, all cards stay in the stack, but cards with lower confidence level appear more often. The task ended after a total of 80 presentations of cards.

The recall task consisted of recalling the energy consumption of an appliance upon appearance of a flashcard with a silhouette on the screen, either by typing numbers in the symbolic condition or by scaling a vertical bar with the mouse in the spatial condition. Thus, recall was conducted in the exact format in which energy consumption was learned. Positive feedback in the form of temporary green highlighting and a checkmark was provided for answers within 10% of the correct value. The task ended when the energy consumption of all appliances was recalled three times. A round of familiarization with the procedure preceded the actual recall task. The recall task was repeated after the comparison task in order to establish whether memory faded, stayed the same, or strengthened as a result of processing.

The comparison task involved determining from memory which of two appliances or sets of three appliances used more energy. First, a central focus point appeared on a computer screen for one second. Then, an item appeared consisting of appliances (their silhouettes) on both the left and the right of the screen. Participants indicated the higher energy consumption by using the keyboard with "Q" for left and "P" for right. One-by-one comparisons involved 28 comparison items, each displaying one appliance on each side. Three-by-three comparisons involved 20 comparison items, each displaying three appliances on each side. The comparisons varied in distance, i.e. the mathematical difference between the magnitudes of the energy consumption values of the appliances. Each comparison item belonged to either one of three distance groups: close, medium, or far (Table 2; see also appendix 9.2 for more detail). Each item was presented four times, twice in its original form (A) and twice mirrored on the screen (B), to compensate for laterality effects. Furthermore, items were presented in a symmetrical sequence (A-B-B-A) in order to balance for learning and fatigue effects.

Table 2. Distances used in the comparison task

	One-by-one		Three-by-three	
	M	*SD*	*M*	*SD*
Close	92.0	34.9	69.9	48.6
Medium	220	36.2	245	18.7
Far	414	77.2	485	148

5.4.3 Procedure

Participants entered a computer lab on campus in groups of up to 12 people and were randomly assigned to conditions. Participants first expressed their informed consent via a computer form, then completed the familiarization task, the learning task, the recall task, the comparison task, and again the recall task. Finally, participants were asked to complete the Form Board Test and the Addition Test, both extracted from the Kit of Factor-Referenced Tests (Ekstrom, French, Harman, & Dermen, 1976a) and adapted into a computer version. After this, participants were thanked and debriefed. The experimental procedure received approval from the department's ethics committee.

5.4.4 Dependent variables

5.4.4.1 First and second recall error

In the beginning and at the end of the study, participants recalled the energy consumption of appliances on the computer, either by typing digits or by dragging a bar with the mouse. The computer automatically calculated how far off the answer was as compared to the correct value. For each recall and for each participant, the absolute difference between the answer (a) and the correct value (b) was divided by the correct value (b), yielding an error percentage ($Error_{\%} = |a - b| / b$). These data were averaged creating a first and a second recall error score for each participant. Inter-item reliability was strong with $\alpha = .815$ in the first recall task and $\alpha = .825$ in the second recall task (together, $\alpha = .895$).

5.4.4.2 Accuracy and response time

Accuracy was the percentage of correctly solved comparison items within a type of comparison and a distance. Each response in the comparison task was recorded as either "correct" (1) or "incorrect" (0). For each comparison type (one-by-one and three-by-three), these data were averaged into three groups of distance (close, medium, and far), creating six accuracy scores per participant, expressed as percentage of correct answers. Across both types of comparisons, a strong inter-item reliability was found, $\alpha = .926$.

Response time (RT) was the delay between an item's presentation and a participant's response, recorded in milliseconds. RTs in the comparison task were screened for outliers. Outliers lie on the tails of a distribution and are caused by other processes than the one under study, such as reflex key presses, fatigue, or loss of attention. Outliers screening followed the method of Cousineau and Chartier (2010) and removed RTs that were either (a) below human perceptual reaction time, taken at 160 milliseconds, or (b) particularly long, with values higher than 2.5 standard deviations above the participant's average RT at this task ($160 \text{ ms} \leq RT \leq M + 2.5\ SD$). This method removed 255 RTs across participants in the symbolic condition and 280 RTs in the spatial condition, which together make up to 3% of the total number of RTs. Only the response times of correct answers were kept in further analysis. For each comparison type (one-by-one and three-by-three) and distance (Close, Medium, and Far) group, the data were averaged creating six mean response times per participant.

5.4.4.3 Numerical and spatial skills

The experiment ended with the completion of two skills tests extracted from the Kit of Factor-Referenced Tests (Ekstrom et al., 1976a) and adapted into a computer version. The first test was the Form Board test, designed to measure spatial visualization, "the ability to manipulate or transform the image of spatial patterns into different arrangements" (Ekstrom, French, Harman, & Dermen, 1976b, p. 173). The test description goes as follows: "Each test item presents 5 shaded drawings of pieces, some or all of which can be put together to form a figure presented in outline form. The task is to indicate which of the pieces, when fitted together, would form the outline" (Ekstrom et al., 1976b, p. 174). The test was presented in two parts each of eight minutes in duration and containing 24 items made out of 5 sub-items, for a total of 240 sub-items. The Spatial Skills score was the number of sub-items marked correctly minus the number of sub-items marked incorrectly. Correctness was directly assessed by the computer. Scores ranged from 36 to 220 ($M = 115$, $SD = 39.2$). Inter-item reliability between sub-items was found at $\alpha = .932$.

The second test was the Addition test, designed to measure number facility, "the ability to perform basic arithmetic operations with speed and accuracy" (Ekstrom et al., 1976b, p. 115). The test was a speed test of mental addition of sets of three 1- or 2- digit numbers. The task was to type in the sum of the numbers in a box below them. The test was presented in two parts, each of 60 items and 2 minutes in length, 120 items total. The Numerical Skills score was the total number of correct answers. Answer correctness was directly assessed by the computer. Scores ranged from 4 to 33 ($M = 17.0$, $SD = 6.39$). Inter-item reliability was found at $\alpha = .736$.

5.5 Results

In order to ensure that only the data from participants who had well succeeded at learning the materials were analyzed, data were filtered according to a dual threshold. Accuracy at the recall and comparison tasks was used as indicator of success at learning the materials. Participants who deviated from the mean by more than one standard deviation in the direction of more error in either task were removed from further analysis. This concerned 11 participants in the symbolic condition and 14 in the spatial condition. Thirty-five participants remained in the symbolic condition, and thirty-eight in the spatial condition ($N = 73$).

5.5.1 Numerical and spatial skills

Numerical skills and spatial skills scores were assessed through the Addition Test and Form Board Test, respectively. Analyses of variance indicated that they did not differ significantly across conditions of external representation, $F_{numerical}$ (1, 71) = 0.45, $p = .506$, $\eta_p^2 = .006$; $F_{spatial}$ (1, 71) = 2.57, $p = .113$, $\eta_p^2 = .035$. Numerical and spatial skills were not significantly correlated with either recall error, comparison accuracy, comparison response time, or one another. Therefore, they were not used as covariates in the analyses.

5.5.2 Recall

In the first completion of the recall task, and after removal from analysis of participants filtered out at the threshold of one standard deviation above the mean as described above, average recall error was found at 10.4% ($SD = 8.98$). This means that, in average, participants recalled a value 10.4% higher or lower than the actual energy consumption of the target appliance. Learning was considered to be sufficient for subsequent execution of the comparison task involving retrieval of the energy consumption values.

In order to determine whether external representation affected recall error and whether recall error changed between the two completions of the recall task, a 2 (time) x 2 (external representation) repeated measures ANOVA was conducted. A significant main effect of external representation was found, $F(1, 71) = 45.13$, $p < .001$, $\eta_p^2 = .389$. Figure 3 shows higher recall when numerical magnitudes were learned with the symbolic external representation. This may be attributed either to the advantage of discreteness for recalling in the symbolic condition or to the difficulty of dragging with the mouse as an input modality in the spatial condition, or to both. Whereas near zero error is expected for copying three digits right after expose, dragging with the mouse to indicate the size of a bar is more cumbersome. A short pilot was conducted to estimate the error for sizing a bar right after exposure. An average 6.66% of error ($SD = 2.99$) was found when sizing a bar with the mouse, explaining part of the 10% difference between the two conditions. Learning was considered sufficient in both conditions with a slight advantage in the symbolic condition.

A significant main effect of time was also found, $F(1, 71) = 15.9$, $p < .001$, $\eta_p^2 = .183$. Figure 3 shows that recall of the appliances' energy consumption generally improved during the comparison task. Repeated access to the energy consumption of the appliances might have strengthened their values in memory.

No significant interaction between external representation and time was found, $F(1, 71) = 0.28$, $p = .601$, $\eta_p^2 = .004$. Learning improved similarly in both symbolic and spatial conditions.

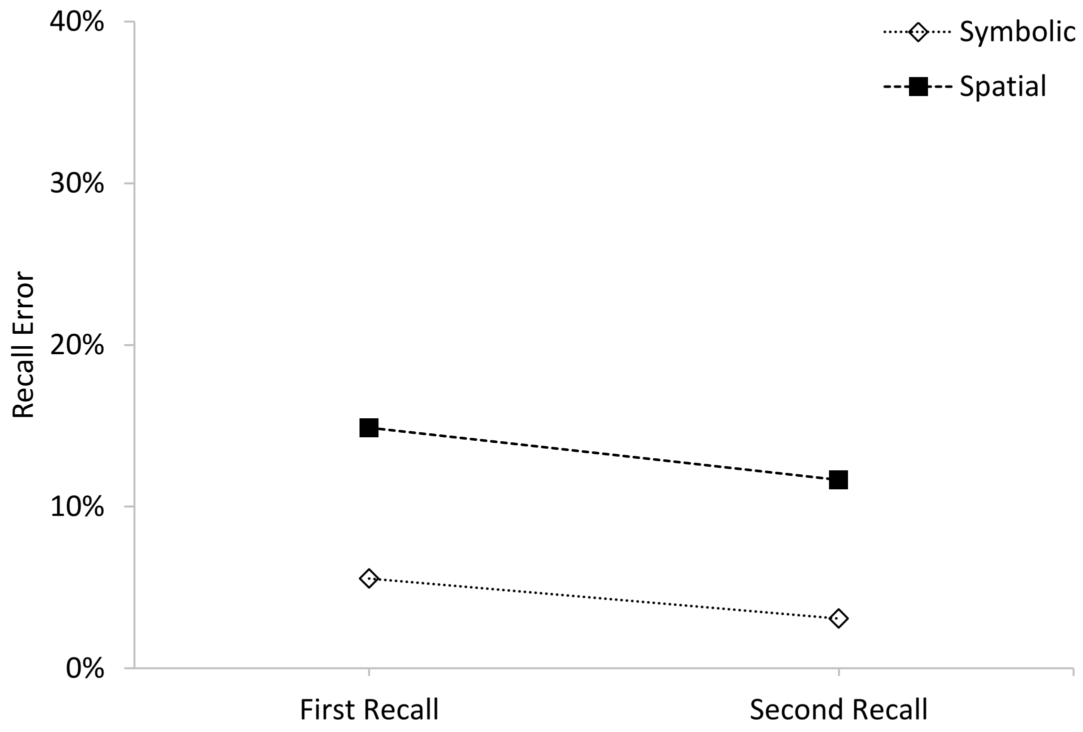


Figure 3. First and second recall error in percentage according to external representation.

5.5.3 Comparisons

According to the spatial congruity of responding hand and larger magnitude, a SNARC effect could have been observed in the comparison task. Specifically, responses could have been faster when the larger magnitude was on the right side of the screen and was thus to be selected with the right hand (see Dehaene et al., 1993). Consequently, response time was analyzed in a 2 (congruity) x 2 (external representation) ANOVA in one-by-one comparisons. No effect of congruity was found, $F(1, 71) = 0.21$, $p = .652$, $\eta_p^2 = .003$. The same analysis was conducted for three-by-three comparisons. In this case as well, no effect of congruity was found, $F(1, 71) = 0.08$, $p = .778$, $\eta_p^2 = .001$. Thus, as hypothesized, the SNARC effect was not observed in this task.

Accuracy and response time were also compared between the two types of comparisons. A repeated-measures ANOVA showed lower accuracy in three-by-three comparisons ($M = 84.6\%$; $SD = 8.11$) as compared to one-by-one comparisons ($M = 91.3\%$; $SD = 5.83$), $F(1,71) = 52.1$, $p < .001$, $\eta_p^2 = .423$. Another repeated-measures ANOVA showed that comparing sets of appliances took significantly longer ($M = 3100$ ms; $SD = 1910$) than comparing single appliances ($M = 1380$ ms; $SD = 436$), $F(1, 71) = 80.4$, $p < .001$, $\eta_p^2 = .538$. These results were expected because three-by-three comparisons involve retrieving and estimating the sum of magnitudes before comparing them, which is a much harder task than mere retrieving and comparing. The much higher complexity of comparing sets of appliances rather than single appliances induces separate statistical analysis.

5.5.3.1 One-by-one comparisons

In one-by-one comparisons, a 3 (distance) x 2 (external representation) repeated-measure MANOVA on accuracy and response time was conducted. The analysis revealed a main effect of distance on accuracy, $F(2, 142) = 72.6$, $p < .001$, $\eta_p^2 = .505$. Figure 4 shows that accuracy was higher when distance was larger, revealing a distance effect. Such a distance effect strongly suggests processing with spatial transient mental representations. There was no significant effect of external representation on accuracy, $F(1, 71) = 0.02$, $p = .885$, $\eta_p^2 < .001$, indicating that no external representation led to more correct responses than the other. No interaction between distance and external representation was found, $F(2, 142) = 0.81$, $p = .448$, $\eta_p^2 = .011$, indicating that the distance effect was equally strong across conditions.

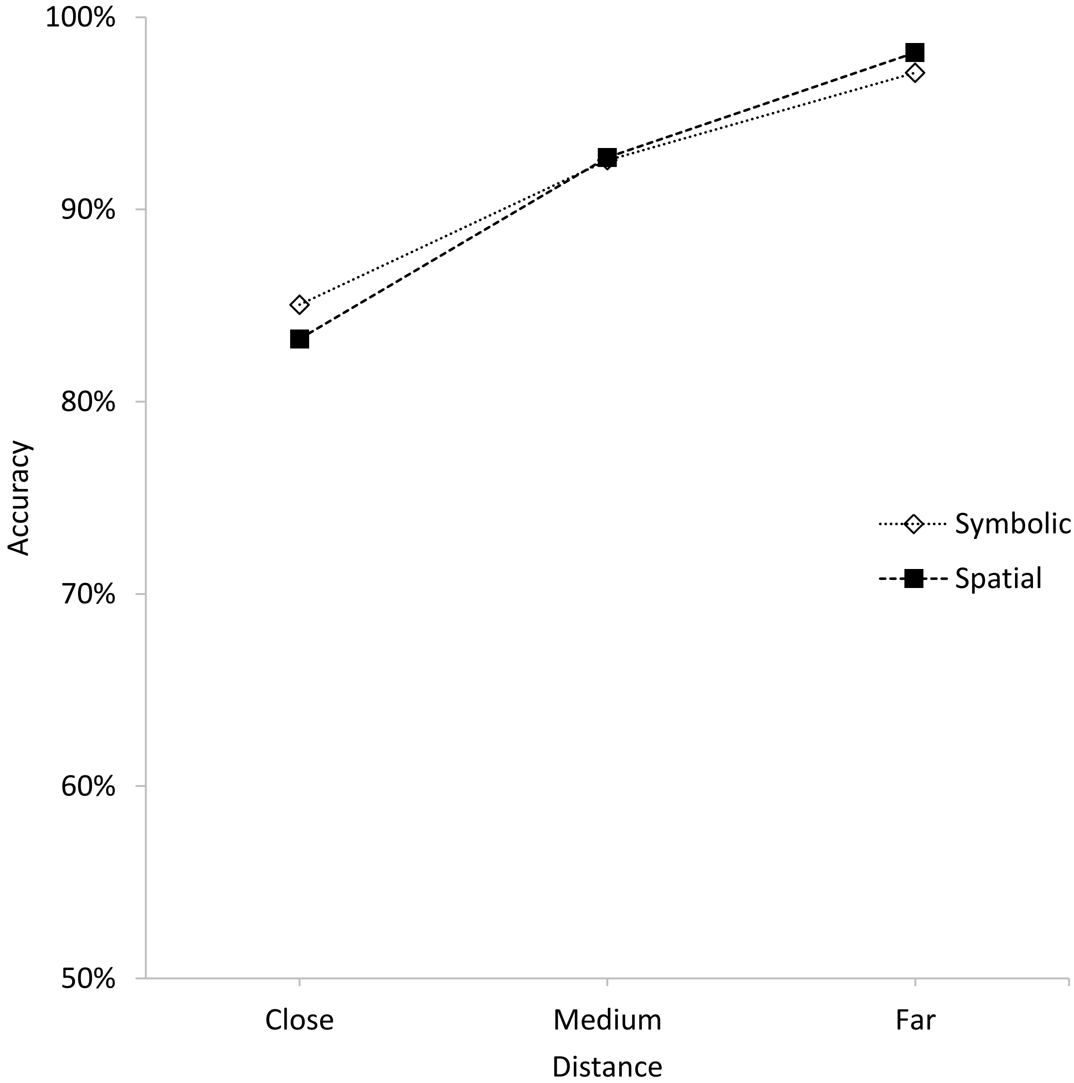


Figure 4. Accuracy for one-by-one comparisons as a function of distance and external representation.

The analysis also revealed a main effect of distance on response time, $F(2,\ 142) = 169$, $p < .001$, $\eta_p^2 = .705$. As figure 5 shows, responses were faster when distance was larger, revealing a distance effect on response time in addition to this found on accuracy. This confirms processing with spatial transient mental representations, consistently with Dehaene's (1992) model in which mental comparisons take place with spatial transient mental representations.

Furthermore, the analysis revealed a main effect of external representation on response time, $F(1,\ 71) = 26.1$, $p < .001$, $\eta_p^2 = .269$, showing significantly faster responses in the spatial condition than in the symbolic condition. This is consistent with the hypothesis that learning magnitudes with spatial external representations allows direct mental comparisons from memory without converting the memorial representation in a different mental format. The longer time required for comparisons in the symbolic condition is attributed to a mental format conversion to obtain a spatial mental representation from the symbolic mental representation stored in memory.

An interaction between distance and external representation was found on response time, $F(2, 142) = 13.4$, $p < .001$, $\eta_p^2 = .159$. The distance effect on response time was stronger in the symbolic condition (Figure 5). This interaction suggests an increased difficulty for comparing similar magnitudes of energy consumption in the symbolic condition. Because this is observed only on response time but not accuracy, the increase in difficulty led to longer mental processes with the same accuracy. A possible explanation could be that in the symbolic conditions, more iterations of a retrieve-convert-compare cycle were needed to obtain sufficient precision in the transient spatial mental representations generated. In other words, the successive steps of 1) retrieval of symbolic mental representation, 2) conversion to an analog representation, and 3) comparison of the sums of energy consumption may or may not actually give a result. In particular, large distances allow for approximate conversions. However, small distances require more accurate conversions for the cycle to enable the comparison. The longer response times for small distances provide evidence for such a retrieve-convert-compare cycle when magnitudes are learned with symbolic external representations. Only when sufficient precision was achieved could the answer be provided, leading to similar accuracy but longer response time.

5.5.3.2 Three-by-three comparisons

In three-by-three comparisons, a 3 (distance) x 2 (external representation) repeated-measures MANOVA on accuracy and response time was conducted. Regarding accuracy, the analysis revealed a main effect of distance as shown in Figure 6, $F(2, 142) = 207$, $p < .001$, $\eta_p^2 = .744$. The distance effect once again supports the hypothesis of spatial magnitude processing in mental comparisons. The effect size seemed to be larger than in one-by-one comparisons. The stronger distance effect may be due to smaller, thus more difficult, distances for Close comparisons and larger, thus easier, distances in Far comparisons in three-by-three comparisons (see Table 2). Accuracy largely decreased for Close distances in three-by-three comparisons.

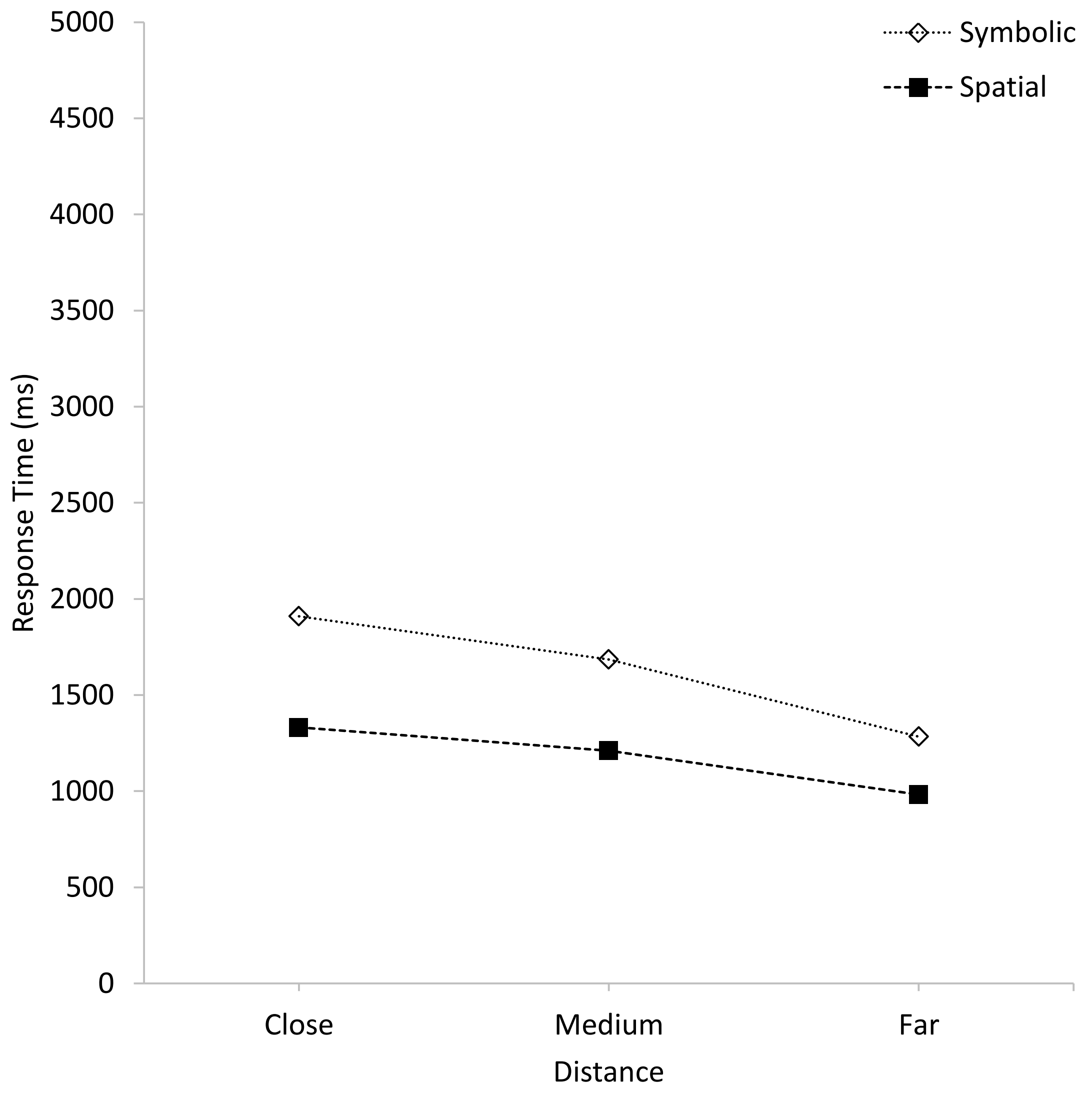


Figure 5. Response time for one-by-one comparisons as a function of distance and external representation.

A main effect of external representation was also found on accuracy, $F(1, 71) = 8.49$, $p = .005$, $\eta_p^2 = .107$, indicating that comparisons were more accurate in the symbolic condition (M = 87.5%; SD = 1.30) than in the spatial condition (M = 82.2%; SD = 1.30). This suggests that learning with a symbolic external representation enabled participants to generate more precise transient spatial mental representations as compared to participants who learned with a spatial external representation. This is also supported by the higher accuracy found on recall in the symbolic condition. Symbolic representations enable more precision. Finally, no interaction between distance and external representation was found on accuracy,

$F(2, 142) = 2.25$, $p = .109$, $\eta_p^2 = .031$, meaning that the distance effect was present to the same extent in both conditions.

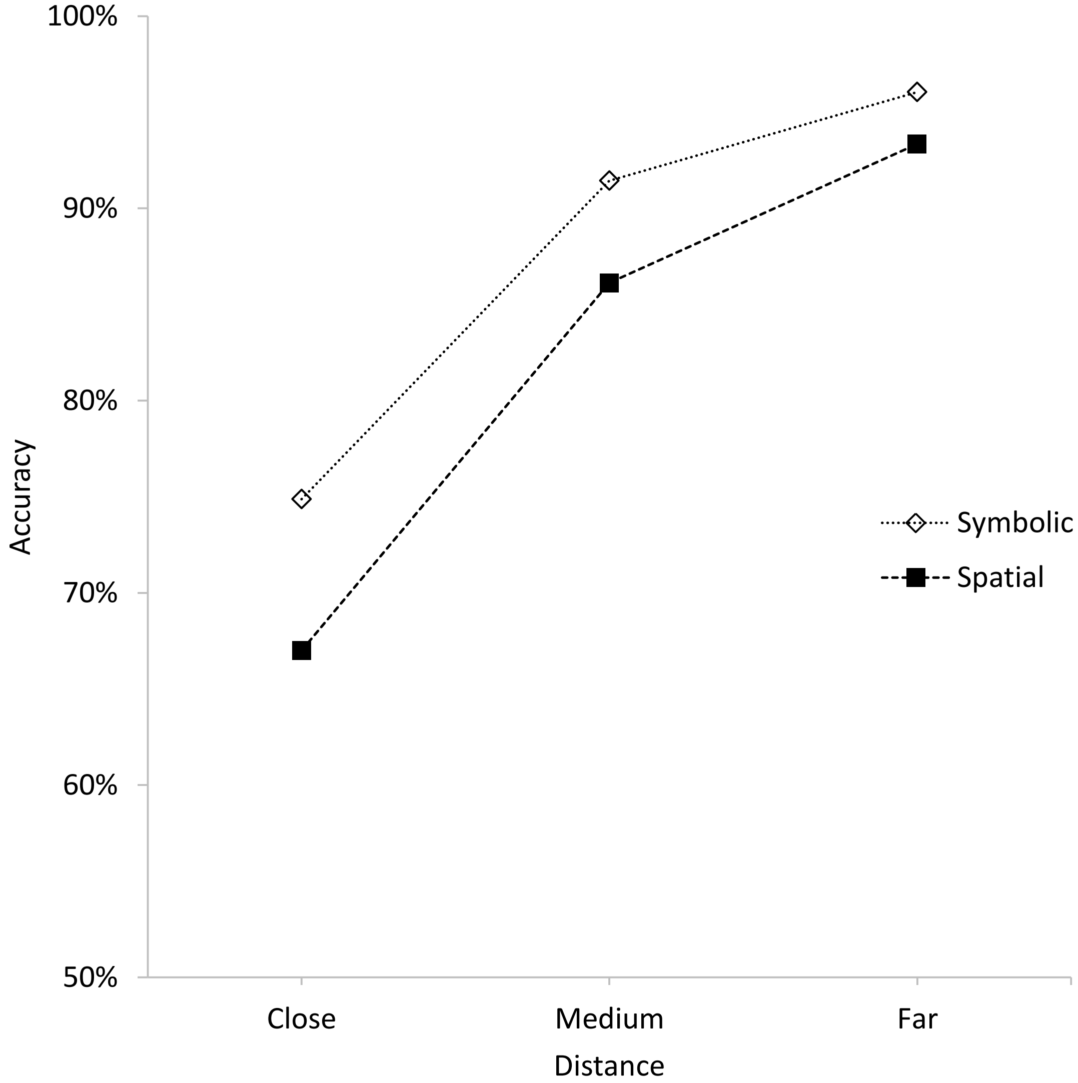


Figure 6. Accuracy for three-by-three comparisons as a function of distance and external representation.

The same analysis also revealed a main effect of distance on response time, $F(2, 142) = 38.5$, $p < .001$, $\eta_p^2 = .352$. This again supports the hypothesis that comparisons are conducted with spatial mental representations. Although the distance effect on response time seems more pronounced in three-by-three comparisons (Figure 7) than in one-by-one comparisons (Figure 5), the effect size is actually smaller. The cause of this probably lies in the nature of the task: the added complexity of processing sums of magnitudes introduced additional variance.

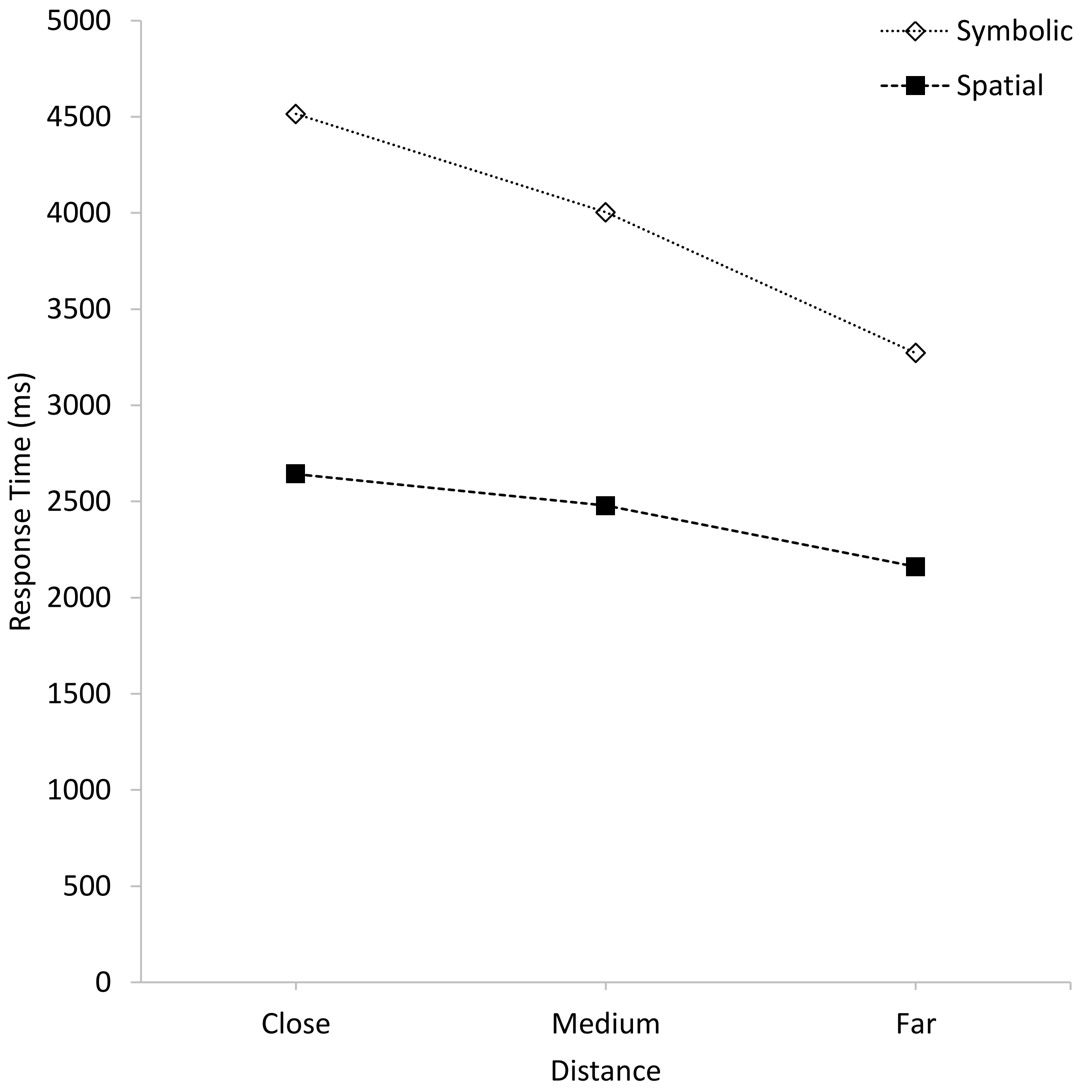


Figure 7. Response time for three-by-three comparisons as a function of distance and external representation.

A main effect of external representation on response time was also found, $F(1,71) = 11.9$, $p = .001$, $\eta_p^2 = .143$, showing that comparing magnitudes of energy consumption was significantly faster in the spatial condition than in the symbolic condition. The effect suggests that comparisons in the symbolic condition required generating a spatial transient representation from the stored symbolic representation. As shown in figure 6, the generation of such a representation leads to higher accuracy but at the cost of speed (Figure 7). Learning with spatial external representations thus leads to much faster yet slightly less accurate comparisons.

Finally, the analysis revealed an interaction between distance and external representation on response time, $F(2, 142) = 7.41$, $p = .001$, $\eta_p^2 = .094$. This interaction indicates that the distance effect on response time was stronger in the symbolic condition (Figure 7). Compared to one-by-one comparisons, three-by-three comparisons require an additional step in the mental mechanism: adding numerical values in order to obtain the total for each set. Again, more iterations of this retrieve-convert-add-compare cycle are required for small distances. In order to obtain high accuracy, this led to longer response times at Close distances (a stronger distance effect).

5.6 Discussion

This study shows that external representation of a magnitude of energy consumption affects later performance at recall and comparison. Results suggest that magnitude is stored in memory as a simulation of the original stimulus, revealing the important role of external representation. A magnitude presented graphically led to higher performance in comparisons and lower performance in recall as compared to the same magnitude presented with digits. The implications of these findings for design and education are discussed in this final section.

5.6.1 Learning and knowing numerical magnitudes

First, results show that learning a magnitude, here a value of energy consumption for fictional household appliances, is possible with both a symbolic and a spatial external representation. This suggests that learning magnitudes from a spatial external representation is feasible and takes place in the same way as learning magnitudes for entities of a spatial nature, stickmen (Kosslyn et al., 1977), circles (Moyer & Bayer, 1976), mentioned in the introduction, or such as learning the relative sizes of animals in everyday life.

As expected, recall error was somewhat lower in the symbolic condition. This effect was attributed to the relative ease of encoding and recalling digits rather than the size of bars. This result could also be interpreted as a sign of higher learning in this condition, especially since accuracy at three-by-three comparisons was found higher in the symbolic condition. However, responses in the comparison task were also significantly faster in the spatial condition than in the symbolic condition. In this respect, the performance with bars can be considered prominent. These results question the notion of knowing a number. Is it important to be able to recall it accurately, or is it important to have a general sense of its magnitude and use it to perform tasks? Since discrete symbols (digits) are commonly used to represent magnitude, it is commonsensical to think in these terms. However, the human mind might have alternative ways depending on the context and the task. Dehaene (1997) describes the number sense as based on a continuous accumulator that does not directly process discrete numbers. The present results support this view. One can have a quite practical sense of magnitude without being able to recall it precisely in digits or to associate a number to it.

5.6.2 Memorial comparisons of magnitudes

According to the literature, comparisons of objects with different magnitudes are conducted with a spatial transient representation. This can be inferred from the observation of a distance effect in comparisons, by which objects with a larger difference in magnitude are compared in shorter time (Moyer & Bayer, 1976). Such a distance effect was found in both accuracy and response time data, in both conditions of external representation (symbolic and spatial), and in both types of comparisons (one-by-one and three-by-three). This effect is a manifestation of the spatial nature of mental comparison processing, where larger differences are more easily decipherable than smaller differences, as is the case in perceptual processes (Henmon, 1906). Presents results suggest that memorial comparisons of magnitudes of energy consumption are always carried out using a spatial transient representation, as shown by Dehaene (1992), irrespective of external representation or complexity of the comparison.

The present study showed that slower comparisons with a symbolic external representation as compared to a spatial external representation. Having learned about magnitudes in digits requires more time in a comparison task. The most parsimonious explanation is that the additional time is taken by the conversion of the symbolic persistent representation into a spatial transient representation. This step is not required when magnitudes have been learned graphically and are stored in a spatial persistent representation. This supports the general hypothesis that persistent mental representations share the properties of initial external representation, as suggested by models of imagery (Kosslyn et al., 1977; Moyer & Bayer, 1976). Furthermore, it shows how transient representations are generated and processed depending on the task (Dehaene, 1992).

The response time results showed that the distance effect was stronger in the symbolic than in the spatial condition for both types of comparisons. Comparing magnitudes took the longest time when comparing sets of one or three appliances with a small difference in magnitude when learned with a symbolic external representation. This may seem counterintuitive as the distance effect is specific to spatial processes, but it suggests that the generation of a sufficiently precise transient spatial representation from a persistent symbolic representation takes even longer when differences are smaller. Several retrieve-convert-compare cycles in simple comparisons, or retrieve-convert-add-compare cycles in complex comparisons might be needed for small differences in reaching the required precision. Also, it may be harder to keep in mind transient spatial representations resulting from a translation than the same transient spatial representations simply recalled from memory. This increased difficulty could be due to a difference in the nature of the persistent representation: symbolic when learning from digits and spatial when learning from a spatial external representation.

Finally, the present results are also consistent with the grounded cognition's mental simulation hypothesis (Barsalou, 1999, 2008; M. Wilson, 2002). The mental process of memorial comparison could be summarized like the following. First, on sight of the object to be compared (or its name/symbol), the associated value is accessed in memory and simulated in the format in which it was originally perceived and encoded (digit/bar). Then, this mental simulation of the value is converted into a spatial mental representation if needed. Finally, the spatial mental representation is mentally compared via visual-spatial cognitive processes as if it were present before the individual's eyes. These analog processes, which make use of existing perceptual mechanisms such as imaging and visual comparison, explain the present results with simplicity and accuracy.

5.6.3 Magnitude sense

The present study results support Dehaene's (1997) viewpoint that a magnitude sense exists independently from numerical notation, and that this magnitude sense is spatial. Learning about magnitude graphically without having access to any Arabic digits led to higher performance on all comparisons, especially the harder ones. Although comparison accuracy was also found to be lower when learning about magnitudes graphically, this nonetheless shows that magnitude understanding does not require numbers in the form of digits for a task of comparison. On the contrary, they may lure people into believing that they are readily able to use a number's magnitude in a task although they don't actually have a clear sense of the magnitude in mind, because they are able to recall the digits. In other words, when a number is learned only through its symbolic representation, the associated magnitude needs to be re-imagined at serious cognitive expense whenever needed.

5.6.4 Practical implications

The first implication is that surface area can be used as a graphical unit, without necessarily needing a digital value always associated to it. Of course, graphical magnitude cannot be expressed verbally, which makes some processes such as communication or perfect recall complicated, but for a single learner and user in energy management tasks, graphical magnitude is actually more practical than digits. Although this may seem obvious to some, home energy management displays are nonetheless covered and cluttered in numbers. The graphical external representations they feature often resemble tables, in which the reader needs to use the graph (e.g., a line) as a cue to find the corresponding value on the axis, as opposed to using the graph as a physical space where size and distance have intrinsic meaning. For example, energy management systems typically display graphs that will scale up or scale down independently on both axes in order to entirely fill the screen. In this scenario, the graphical unit, i.e., pixel or inch, loses any meaning, and the only way to infer magnitude is to read it from the axes. Another common external representation is this found for energy efficiency on the European Union energy label. On this label, different bar lengths correspond to different levels of energy efficiency, with longer bars corresponding to higher energy consumption. However, the graphical units (e.g., centimeters) do not correspond to any magnitude, they are arbitrary and serve only as an ordinal display. Next to this graphical external representation is a digital number: the annual estimated energy consumption of the appliance. This important information, however, is not represented graphically on the label. Our results suggest that a graphical unit of energy consumption could be successfully used, if it featured spatial consistency.

Although our results indicate that digits increase one's ability to accurately recall numbers, magnitude sense is what numerical external representations actually seek to convey. Digits may create the illusion of knowledge and understanding, as it happens in classrooms. Dehaene (1997) describes that magnitude sense is too often neglected in children's education, their ability to perform calculations being the only ruler by which their mathematical skill is measured. In consequence, instead of developing a good sense of magnitude, children are mostly trained to compute values to which they often don't associate meaning. Similarly, energy users should not be thought of as information-processing systems computing values and making rational decisions based on a pre-established flowchart. In fact, smart home energy management systems computers fail to replace human control so far. People are needed in energy management for their quick and efficient heuristic comparison skills based on fuzzy spatial mental representations. Spatial external representations should therefore be favored in energy management and other domains.

5.6.5 Limitations

The limitations of this study principally reside in the possibility for generalizing the findings to real world settings. Designing a learning and a comparison tasks corresponding to home energy management while maintaining control over all variables led to compromises that could be perceived as shortcomings. First, a bias from prior knowledge of existing appliances was avoided by designing imaginary appliances. In real life, citizens may have a certain sense of the energy consumption associated to each appliance. Also, due to the use of a criterion of accuracy on the experimental tasks, results cannot be generalized to people who haven't sufficiently learned and shown ability to recall and compare magnitudes. In addition, different instructions in the learning phase, such as the instruction to visualize a verbal description in Denis' (2008) study, may lead to different persistent mental representations and thus different experimental results. Furthermore, the delay between the first and the second recall was only about half an hour in the present study. Finally, there is always a trade-off between accuracy and response time. The instructions here were to aim for the maximum possible number of correct responses per minute. In an actual everyday setting, either accuracy or speed could be more important. Future research could take into account these variables related to prior knowledge of the appliance, the delay between learning and using values of energy consumption, and the particular value of the trade-off between speed and accuracy.

5.6.6 Future research

Many possible analog external representations could be used to represent energy consumption or other numerical magnitudes. Future research should test other external representations in order to determine whether their effects vary on different tasks and which variables would cause the difference. The strong link between space and magnitude suggests that analog visuospatial external representations are ideal for comparison, but other modalities could be explored as well, such as pitch height or size as perceived by the haptic sense.

Moreover, it is possible that the combination of multiple analog external representations would further ease mental comparison. Larger magnitudes could for instance be represented on a graph with both a longer bar and a warmer color, combining the effects of both analog external representations.

Finally, the most appropriate external representation actually depends on the task to be conducted. In the present study, a symbolic external representation, Arabic numerals, was preferable for recall, and a spatial external representation, graphical bars, was arguably more appropriate for comparisons. Future research could include still other essentially different instructions at the time of learning or essentially different tasks.

6

LEARNING AND COMPARING ENERGY CONSUMPTION: EFFECTS OF SPATIALITY, GROUNDEDNESS, AND PHYSICALITY OF EXTERNAL REPRESENTATIONS [2]

Energy management requires learning and comparing magnitudes of energy consumption. Energy management tasks typically consist of comparing the energy consumption of appliances or of households, between one another or against some reference. Such mentally comparisons involve a specific domain of numerical cognition, the processing of magnitude. Magnitude processing can be influenced by the external representations of magnitude provided to a learner.

6.1 Processing numerical magnitude

According to Dehaene's *triple-code model* (1992), numbers are processed in three different kinds of mental representations, corresponding to real-life formats. The first kind of representation is the visual Arabic representation, which allows the processing and manipulation of numerals in calculus. Multiplying a number by 10 by simply adding a zero at its tail relies solely on this representation. The second kind of representation is the auditory verbal word representation, which allows number words to be spoken and understood verbally

[2] This chapter has been written as a manuscript (Galilée, de Vries, & Scheiter, in preparation) to be submitted for publication in a journal. The study was conducted at the Knowledge Media Research Center, IWM Leibnitz Institut für Wissenmedien, Tübingen, Germany. The stay was funded by a CMIRA Explo'RA Doc scholarship. Many thanks to Katharina Scheiter for enabling this research project.

in a specific language. Rote learned multiplication tables use this representation. Finally, the third kind of representation is the representation of magnitude, or 'amount of stuff'. Magnitude representation is based on an analog mental representation with spatial properties. Typically, it takes the form of a mental number line, a tool used in mathematics where distance (differences in space) represents differences in magnitude. Larger magnitudes are represented further to the right than smaller magnitudes; larger differences in magnitudes are reflected in larger distances on the number line. Operations on magnitude such as estimations and comparisons rely on this external representation.

Because numerical comparisons are conducted with analog mental representations, the comparison of two numbers that are distant on the number line is easier and faster than the comparison of numbers that are close to one another (Moyer & Landauer, 1967). In other words, the comparison of the pair [2, 9] is faster than the comparison of the pair [2, 3]. This *distance effect* (Moyer & Bayer, 1976) is expected in all numerical comparisons and is an evidence of analog processes at use.

When numerical tasks are conducted from memory, mental representations are generated from memory. The way a magnitude is initially learned, for instance, with symbols such as Arabic numerals or with a spatial external representation such as a bar, affects performance at tasks conducted from memory (see chapter 5). Generating mental representations from memory is achieved via mental simulations grounded in perceptual processes (Barsalou, 2008; M. Wilson, 2002). Mental representations consequently share the properties of external representation because they use the same perceptual processes. Thus, a learner's performance on different tasks of magnitude processing will vary with external representation of magnitude.

6.2 External representations of magnitude

In the Western culture, magnitude is often conveyed with Arabic numerals, a symbolic external representation. Symbolic external representations are characterized by arbitrary relationships between representation and what is being represented (object). Meaning is achieved through convention and cultural agreement. However, spatial external representations also present advantageous alternatives according to task demands. A growing body of literature also focuses on tangible external representations which feature, among others, two separate dimensions that are often conflated: groundedness, and physicality

(Marshall, 2007). In the following, the possible advantages of spatial, grounded, and physical external representations for tasks related to numerical magnitudes are discussed.

6.2.1 Symbolic external representations

In Peirce's definition (Peirce, 1906), *symbols* are signs (i.e., external representations) that represent by virtue of convention and *icons* are signs that represent by partaking in the characters of the represented object. Symbolic external representations bear no resemblance with what they represent. For instance, the numeral "7" is not larger than the numeral "2", rather it has been defined that "7" represents an amount of larger magnitude. On the contrary, iconic external representations have with their object a character in common. Magnitude is borne by iconic external representations on a dimension such as color intensity, number of dots, or typically amount of space covered by ink or pixels in a graph. Other examples include coins or tokens representing money, or space between a fisher's hands signifying the size of a catch.

6.2.2 Spatial external representations

The most adequate dimension on which to represent magnitude in an iconic external representation seems to be space because mental magnitude comparisons are conducted on a spatial mental representation. According to the triple-code model (Dehaene, 1992), magnitude is processed on an analog horizontal number line, that is, a spatial mental representation. Other spatial mental representations than a number line can also be used in magnitude representation such as a clock face (D. Bächtold et al., 1998), or a vertical line or bar (Rusconi et al., 2006, see also chapter 5). The link between space and magnitude is strong, as shown by spatial biases of numerical processing named Spatial-Numerical Associations (SNAs) (M. H. Fischer & Brugger, 2011). For instance, the Spatial-Numerical Association of Response Codes (SNARC) in parity judgments induces faster responses to stimuli of small magnitude with the left hand and faster responses to stimuli of larger magnitude with the right hand (Dehaene et al., 1993) and was abundantly replicated (Wood et al., 2008). Other findings support the importance of spatial mental processes in magnitude processing and representation (for a review, see de Hevia et al., 2008).

Spatial external representations are however not well adapted for precise recall, because they do not provide the precision that symbols such as numerals allow. As a consequence, spatial external representations, as compared to symbolic external representations, have been shown to lead to more error at a recall task (chapter 5).

6.2.3 Abstract and grounded external representations

Representations can be grounded in two ways: contextually grounded and perceptually grounded (Braithwaite & Goldstone, 2013). Contextually grounded representations, e.g. descriptions of realistic situations, can improve performance at visual-spatial tasks such as mental magnitude comparisons. Contextually grounded representations are representations relying on a situated, rather than abstract, context. According to a meta-analysis of neuro-imaging studies, grounded concepts engage the perceptual system to a greater extent as compared to abstract concepts which engage more the verbal system (Wang, Conder, Blitzer, & Shinkareva, 2010). Grounded concepts could thus provide support for spatial mental representations of magnitude. Moreover, contextually grounded representations may ease cognitive tasks according to the *situation facilitation hypothesis*, according to which context can cue relevant knowledge facilitating effective strategy selection and execution (Koedinger & Nathan, 2004). There is mixed evidence that appropriate context improves problem solving performance in mathematics problems (Baranes, Perry, & Stigler, 1989; Koedinger & Nathan, 2004).

Evidence is mixed about the effects of perceptually grounded external representations such as photos or realistic drawings. Perceptually grounded external representations are realistic and perceptually rich representations of the object they represent and may be defined as depictions of the object (see Schnotz & Bannert, 2003). Conversely, representations that are not grounded are described as abstract, idealized (Goldstone & Son, 2005), schematic (Scheiter, Gerjets, Huk, Imhof, & Kammerer, 2009), or formal (Braithwaite & Goldstone, 2013). Perceptually grounded external representations improve learning for students with prior knowledge of a topic (Joseph & Dwyer, 1984) or when used in *progressive formalization* or *concreteness fading*, a scaffolding strategy in which perceptually grounded external representations are provided to learners before abstract representations (Goldstone & Son, 2005). Perceptually grounded external representations of electrical circuits parts, such as drawings of lightbulbs, have also been shown to increase problem solving performance when mixed with abstract symbols (Moreno, Ozogul, & Reisslein, 2011). Such improvement, however, cannot always be observed. Scheiter and colleagues (2009) found that learning of the process of mitosis was poorer when supported only by perceptually grounded external representations, as compared to any other combination of perceptually grounded and symbolic external representations. The value of perceptually grounded external representations thus resides in their combination with other representations for progressive formalization. In the case of magnitude processing, however, formalization of representation may not be a major obstacle. Consequently, perceptually grounded external representations may rather be useful in supporting contextual grounding, the details of representations evoking the context. In other words, a certain level of realism in external representations may provide context, but further realism would not affect magnitude processing.

6.2.4 Physical external representations

Physical representations may ease mental comparisons by convergence of the multiple external iconic representations they intrinsically contain. Representations featuring physical parts that can be manipulated are called physical representations. These physical representations are not the objects they represent. For instance, physical objects named manipulatives are used in math education (e.g., Ball, 1992; Sarama & Clements, 2016; Uttal, Scudder, & DeLoache, 1997) and are merely representations of mathematical concepts. The most important property of physical representation is that they intrinsically contain multiple external representations (Marshall, 2007). For instance, blocks of different sizes representing different amounts of energy would represent by virtue of their visual size, of the width of the hand grip necessary to grasp them, and of their weight—three iconic external representations in themselves. Similar graphical representations would only represent by virtue of visual size. Multiple external representations can enhance learning according to their characteristics in reference to their function and the related cognitive tasks (Ainsworth, 2006; Cheng, 1999; Lund & Bécu-Robinault, 2010; Scheiter et al., 2009). However, according to Marshall (2007), research has so far failed to show that physical representations are better suited than graphical representations for cognitive tasks and learning. Some studies found mixed evidence (Klahr, Triona, & Williams, 2007; Uttal et al., 1997) or absence of a significant difference (Fails et al., 2005).

6.3 Current study

The present study investigated the effects of four different external representations of energy consumption on performance at both recall and magnitude mental comparisons conducted from memory. In the first experimental condition, the external representation of energy consumption was symbolic and consisted of Arabic numerals. In the second experimental condition, the external representation of energy consumption was spatial (and thus iconic), and consisted of a line of colored dots on paper. In the third experimental condition, the external representation of energy consumption was spatial and perceptually grounded because it consisted of a line of drawings of lightbulbs rather than dots. The representation was also contextually grounded because the unit of magnitude was described in reference to the energy use of real-world lightbulbs, with one symbolic lightbulb representing the energy consumption of one real-world lightbulb whereas in abstract external representations, no real-word equivalent for the energy unit was provided. Finally, in the fourth experimental condition, the external representation of energy consumption was spatial, grounded, and physical, and consisted of a line of lit-up actual lightbulbs on an apparatus. Like the spatial external representations, it intrinsically possessed (1) a spatial dimension. However, it possessed as well (2) a dimension of luminosity, which increased with the number of lightbulbs lit-up, (3) a dimension of heat, because the lightbulbs radiated infrared light that could be sensed on the skin, (4) a sound duration due to the number of switches clicking when the lightbulbs were turned on, and (5) a dimension of manual effort, also linked to the number of switches to turn on by the participant (up to 19 switches at a time). In consequence, this makes the lightbulb-based representation an example of a physical representation. The energy flow could not actually be felt by the haptic sense, however it was carried by the luminous and heat fluxes, and the manual interaction on the light switches provided relevant physical contact.

Regarding procedure, a first phase consisted of learning about appliances and their energy consumption from either one of the four external representations. A second phase consisted of recalling the energy consumption from memory. This was conducted with the same external representation as used for learning. A third and final phase consisted of comparing the magnitude of energy consumption of different appliances from memory. In a first type of mental comparisons, the energy consumption of a single appliance was compared to a single other one (one-by-one comparisons). This involved retrieval of the magnitude of energy consumption, then the generation of a spatial mental representation, and finally the mental comparison. In a second and more complex type of mental comparisons, two pairs of appliances were compared to one another (two-by-two comparisons), which additionally involved the mental manipulation of more magnitudes as well as an estimation of sums of magnitude. In both types of comparisons, the difference in magnitudes to be compared varied. This difference is called *distance*, in reference to the *distance effect* stipulating that larger differences in magnitude are faster to compare (Moyer & Landauer, 1967). External representation and distance between to-be-compared appliances in the comparison phase were taken as independent variables. Dependent variables were recall accuracy, comparison accuracy, and comparison response time. Because one-by-one and two-by-two comparisons were variations of a similar task, their data were to be analyzed similarly but separately.

6.3.1 Hypotheses

Two main hypotheses ensued from theory. A distance effect was expected, as well as an effect of external representation.

H1: Distance effect. Due to the spatial processing of magnitude with the analog magnitude representation (Dehaene, 1992), a distance effect was expected in comparisons, in the form of shorter response time for comparisons featuring larger distances. A distance effect is also expected on accuracy, with higher accuracy for larger distances, as found in the previous study (chapter 5).

H2. External representation effect. (a) Recall accuracy was expected to be higher with the symbolic external representation than with any spatial external representation because symbols are discrete and allow precise recall. On the contrary, although the spatial external representations used in this study consisted of a discrete number of dots, the line of dots could be processed as a spatial continuum which does not have such precision. (b) Comparisons were expected to be faster with spatial external representations, as compared to symbolic external representations, due to the availability of an adapted (spatial) mental representation for comparisons. (c) Comparisons were expected to be faster and more accurate with grounded external representations as compared to abstract representations, due to the contribution of contextual cues in the generation of mental representations. (d) Comparisons were expected to be faster and more accurate with the physical external representation as compared to spatial external representations due to the encoding of representations in multiple modalities converging in a more accurate representation and facilitating comparisons.

Further, the presence of a SNARC effect was considered but hypothesized. Due to the spatial orientation of the number line, horizontal from left to right, a Spatial-Numerical Association of Response Codes (SNARC) effect is often observed in parity judgments and consist of faster responses to stimuli of small magnitude with the left hand and faster responses to stimuli of larger magnitude with the right hand (Dehaene et al., 1993). If participants used the number line as mental representation of magnitude in present, the SNARC effect could be observed. However, comparisons is the present study being conducted from memory, response times were expected to be longer than in studies targeting the SNARC effect, resulting in a lack of sensitivity to detect the mild SNARC effect. Consequently, the SNARC effect was not expected.

6.4 Method

6.4.1 Participants

Ninety-three undergraduate students (66 female) from various fields were recruited at a German University and rewarded with either money or extra credits. Participants were randomly assigned to one of the four experimental conditions described below, with 24 participants serving in condition (a), 24 participants serving in condition (b), 23 participants serving in condition (c), 22 participants serving in condition (d). Age ranged from 18 to 35 years with a mean age of 22.4 years (*SD* = 3.39).

6.4.2 Independent variables

Two independent variables were manipulated: external representation and distance. Firstly, external representation was manipulated as between-participant independent variable, leading to four experimental conditions. Energy consumption was represented by either a (a) symbolic abstract external representation (with Arabic numerals), (b) spatial abstract external representation (a line of dots drawn on paper), (c) spatial grounded external representation (a line of lightbulbs drawn on paper), or (d) physical grounded external representation (a row of lit-up actual lightbulbs on an apparatus). The second independent variable, distance, referred to the numerical difference between energy consumption of appliances in the comparison task and was varied within-participants.

6.4.3 Materials and apparatus

Three tasks were extracted from a previous study (chapter 5) and adapted for the present experiment: (1) a learning task, where participants learned about the energy use of appliances, (2) a recall task about this energy use, and (3) a task of mental comparison of energy use.

Name: Rainbow Maker
Used in: Kid's Room
The Rainbow Maker makes crayons of any color imaginable, with or without glitter.

Figure 8. An example of an appliance (translated to German in the study).

6.4.3.1 Learning task

The learning task concerned six fictitious electric appliances defined by their silhouette, name, short textual description, and energy consumption. The silhouettes were carefully drawn to be equally different from one another, so that no silhouette was more recognizable than the others. All silhouettes were in black and white and contained the same number of black pixels within one percent. The names and descriptions were chosen not to evoke a particularly low or high energy use (Figure 8; see also appendix 9.1).

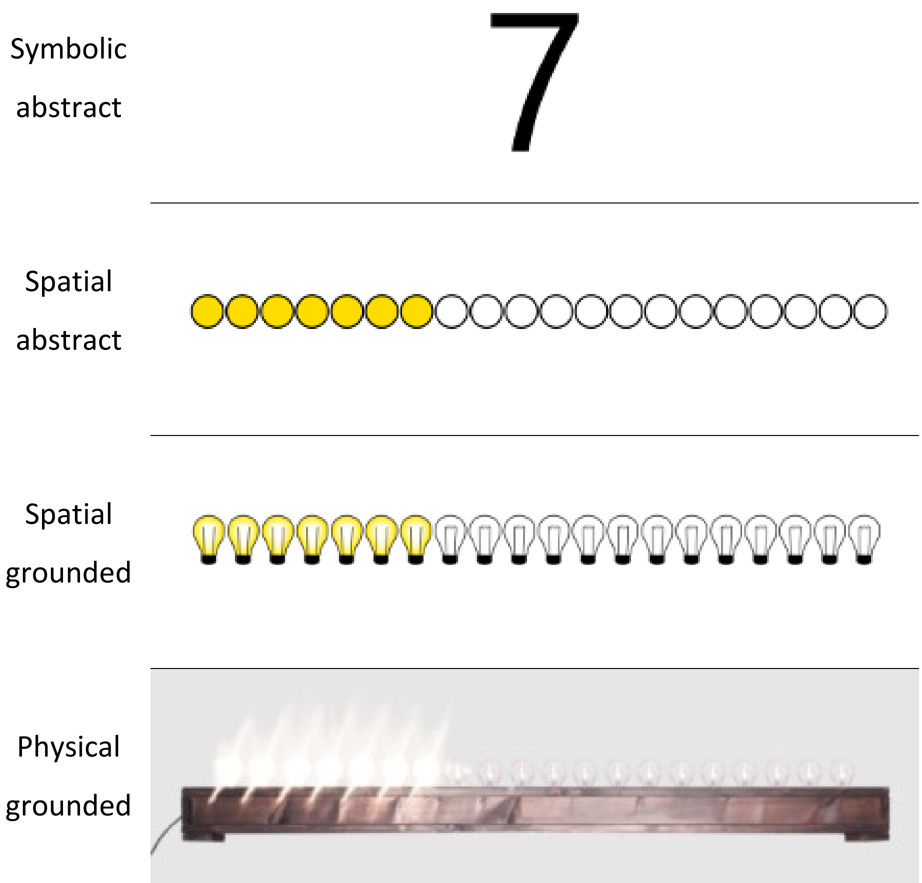


Figure 9. The four external representations.

Appliances' silhouettes were presented one at a time by the experimenter. On the participant's request, the corresponding energy use was shown in one of the four types of representations: (1) as black digits on paper, in the symbolic abstract condition, (2) as yellow-filled dots in a row of 20 dot outlines on paper, in the spatial abstract condition (3) as yellow-filled lightbulbs in a row of 20 lightbulb pictures on paper, in the spatial grounded condition, or (4) as lit-up real lightbulbs out of the row of 20 on an experimental apparatus, in the physical grounded condition (Figure 9). The energy use values were in increments of 3 units, ranging from 4 to 19 units. All appliances were presented six times in random order.

An experimental device similar to a lamp was crafted for the purpose of showing the energy use of the appliances in the physical grounded condition. The device consisted of a horizontal 150cm-long wooden beam carrying a row of 20 lightbulbs on the top, connected in parallel, and which each had an individual switch below. The lightbulbs were typical home lightbulbs, eco-halogen, with an output of 370 lumens at 2800 kelvins.

6.4.3.2 Recall task

The recall task consisted of recalling the energy use of an appliance upon presentation of its silhouette. To respond to a recall item, participants had to write down numbers in the symbolic abstract condition, color dots or lightbulb pictures in the spatial conditions, or turn on the right number of lightbulbs on the apparatus in the physical condition. In order to provide feedback, any error was corrected by the experimenter. The task was preceded by a round of familiarization with the procedure, and ended when the energy use of all appliances had been recalled three times, correctly or not. The percentage of erroneous recalls varied from none to 61.1% ($M = 9.50$, $SD = 14.1$, $\alpha = .817$) and served as dependent variable.

6.4.3.3 Comparison task

The comparison task was conducted on a computer. It involved determining from memory which of two appliances (in one-by-one comparisons) or pairs of appliances (in two-by-two comparisons) used more energy. First, a central focus point appeared on a computer screen for one second. Then, silhouettes of appliances appeared on both left and right of the screen: a single silhouette on both sides in one-by-one comparisons, and a pair of silhouettes on both sides in two-by-two comparisons. Participants indicated the side of higher energy use by pressing a button either on the left or on the right of a response pad. The comparisons varied in distance, that is, the numerical difference between the energy use values of the appliances, to create a within-subject variable. In one-by-one comparisons, five distances were presented, ranging from 3 to 15 units of energy in increments of 3. In two-by-two comparisons, three distances were presented: 3, 6, and 9 units of energy (see appendix 9.3 for detailed distances in both types of comparisons). Both types of comparisons contained 15 different comparison items each. Each item was presented four times, twice in its original form (A) and twice mirrored on the screen (B), to compensate for laterality effects. Furthermore, the items were presented in a symmetrical pattern (A-B-B-A) in order to counterbalance learning and fatigue effects.

Each response in the comparison task was recorded by the computer as either correct (1) or incorrect (0). For each comparison type (one-by-one and two-by-two), these data were averaged into either five or three groups of distance (respectively) in order to enable testing the distance effect, and recorded as accuracy scores, expressed as percentage of correct answers out of the 60 comparisons of each type.

The delay between an item's presentation and a participant's response, the response time, was recorded in milliseconds. For each comparison type (one-by-one and two-by-two), these data were averaged into the same groups of distance as used for accuracy scores, creating eight scores of response time per participant, expressed in milliseconds. Response times (RTs) in the comparison task were screened for outliers. Outliers are scores that lie on the tails of a distribution and are caused by other processes than the one under study, such as reflex keypresses, fatigue, or loss of attention. Some other scores may also be caused by other processes than the one under study, but falling close to the participant's average score, they can never be identified as outliers (Ratcliff, 1993). Outliers screening followed the method of Cousineau and Chartier (2010) and removed RTs that were either (a) below human perceptual reaction time, taken at 160 ms, or (b) particularly long, with values higher than 2.5 standard deviations above the participant's average RT at this task ($160 \text{ ms} \leq RT \leq M + 2.5\ SD$). This method removed 108 RTs across participants in the symbolic abstract condition, 123 in the spatial abstract condition, 113 in the spatial grounded condition, and 111 in the physical grounded condition, which together make up 8.01% of the total number of RTs. Only the response times of correct answers were kept for further analysis.

6.4.4 Dependent measures

Dependent variables were recall accuracy, comparison accuracy, and comparison response time. Recall accuracy was the percentage of erroneous recalls at the recall task. Comparison accuracy was the percentage of correctly solved comparisons. Comparison response time was the delay in milliseconds between item presentation and participant's response.

6.4.5 Procedure

Each participant entered individually an office on campus and was randomly assigned to an experimental condition. The participant first expressed her informed consent via a computer form. Then, she was presented with the names, descriptions, and silhouettes, but not the energy use, of the six appliances. She was instructed to remember as much about them as possible and asked questions about them. She then completed the learning task, the recall task, and the comparison task. There were no time limits for any of the tasks. After this, the participant was thanked and debriefed. A session lasted approximately 45 minutes. The experimental procedure received approval from the department's ethics committee.

6.5 Results

In order to ensure that only the data from participants who had well succeeded at learning the materials were analyzed, data were filtered according to a dual threshold. Accuracy at the recall and comparison tasks was used as indicator of success at learning the materials. Participants who deviated from the mean by more than two standard deviations in the direction of more error in either task were removed from further analysis. This concerned 1 participant in the symbolic abstract (a) condition, 2 in the spatial abstract (b) condition, 2 in the spatial grounded (c) condition, and 4 in the physical grounded (d) condition. Accordingly, 23 participants remained in the symbolic abstract (a) condition, 22 in the spatial abstract (b) condition, 21 in the spatial grounded (c) condition and 18 in the physical grounded (d) condition ($N = 84$).

6.5.1 Recall

An ANOVA was conducted to determine the effect of external representation on percentage of erroneous recalls at the recall task, using planned repeated contrasts to compare (1) the symbolic abstract and the spatial abstract conditions, (2) the spatial abstract and the spatial grounded conditions, and (3) the spatial grounded and the physical grounded conditions. The analysis showed a significant effect of external representation, $F(3, 80) = 5.38$, $p = .002$, $\eta_p^2 = .168$, where only the contrast between the symbolic abstract and the spatial abstract conditions showed a significant difference ($p = .007$). Means show that fewer errors were made in the symbolic abstract condition than in the spatial abstract condition (Figure 10). A t-test further indicated that the percentage of erroneous recall in the symbolic abstract condition did not significantly differ from zero, $t(22) = 1.45$, $p = .162$. This supports the hypothesis (H2a) that a symbolic representation leads to less erroneous recall as compared to a spatial or physical external representation. A symbolic representation seemed to be most appropriate for the recall task, with nearly zero error.

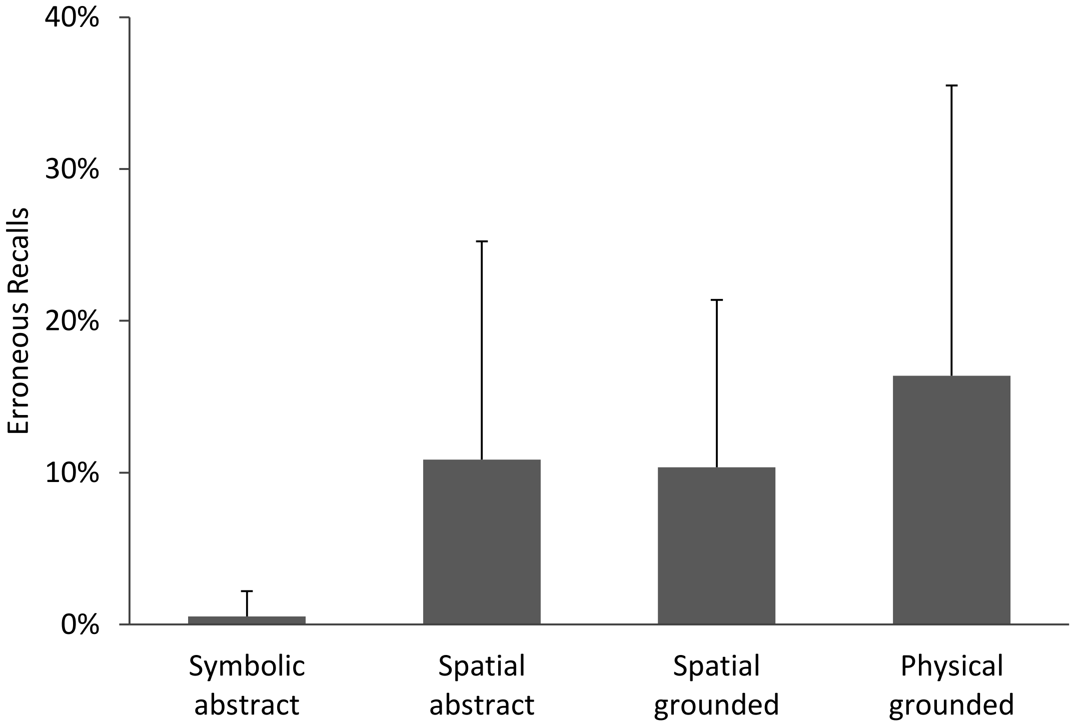


Figure 10. Percentage of erroneous recalls per condition, with standard deviation.

6.5.2 Comparisons

According to the spatial congruity of responding hand and larger magnitude, a SNARC effect could have been observed in the comparison task. Specifically, responses could have been faster when the larger magnitude was on the right side of the screen and was thus to be selected with the right hand (see Dehaene et al., 1993). Response time was analyzed in a 2 (congruity) x 2 (external representation) ANOVA in one-by-one comparisons. No effect of congruity was found, $F(1, 80) = 2.40$, $p = .125$, $\eta_p^2 = .029$. The same analysis was conducted for two-by-two comparisons. In this case as well, no effect of congruity was found, $F(1, 78) = 0.65$, $p = .422$, $\eta_p^2 = .008$. Thus, the SNARC effect was not observed.

Then were analyzed the effects of external representation and distance on accuracy and response time. All analyses were conducted with planned repeated contrasts to compare the effect of external representation between (1) the symbolic abstract and the spatial abstract conditions, testing spatiality, (2) the spatial abstract and the spatial grounded conditions, testing groundedness, and (3) the spatial grounded and the physical grounded conditions, testing physicality.

6.5.2.1 One-by-one comparisons

In one-by-one comparisons, a 5 (distance) x 4 (external representation) repeated-measures MANOVA on accuracy and response time was conducted. The MANOVA revealed a main effect of distance on accuracy, $F(4, 320) = 57.7$, $p < .001$, $\eta_p^2 = .419$. As figure 11 shows, this distance effect goes in the expected direction: comparisons across a greater distance were conducted with higher accuracy, supporting H1.

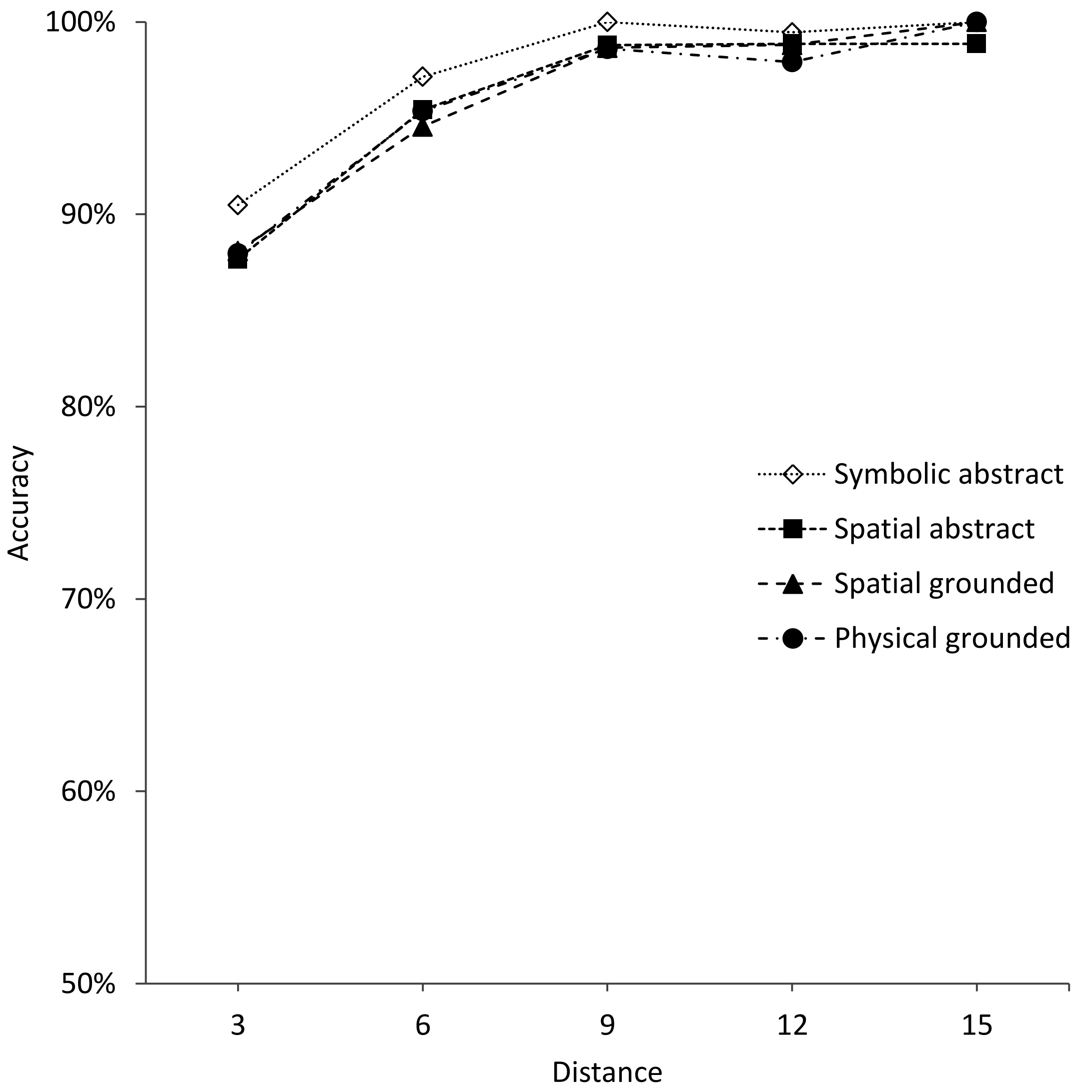


Figure 11. Accuracy for one-by-one comparisons as a function of distance and external representation.

The presence of a distance effect is evidence that comparisons were conducted with spatial mental representations. Moreover, a ceiling effect can be observed in figure 11. Comparisons at large distances all reached very high accuracy close to 100%, suggesting that these comparisons were too easy to allow the detection of any effect of distance or condition. This indicates that more difficult tasks should be used in future studies. External representation had no effect on accuracy, providing no support for H2 in this analysis, and there was no interaction between external representation and distance on accuracy ($F < 1.5$ in both cases). No differences in accuracy between conditions of external representation are observable (Figure 11).

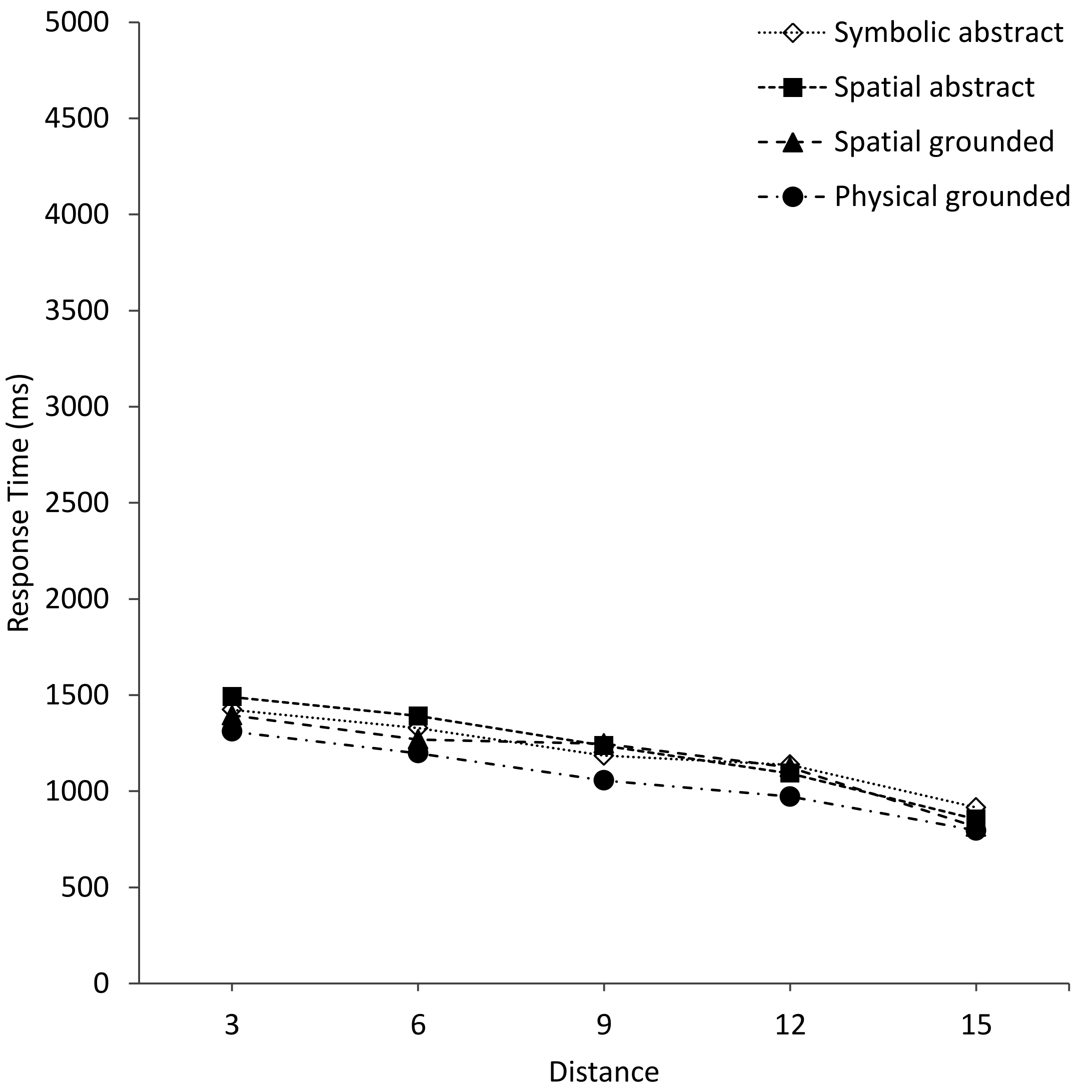


Figure 12. Response time for one-by-one comparisons as a function of distance and external representation.

The analysis also revealed an effect of distance on response time, $F(4, 320) = 113$, $p < .001$, $\eta_p^2 = .587$ (Figure 12). Comparisons across a greater distance were easier, as reflected in both higher accuracy and shorter response times. The distance effect again indicates the processing of comparisons with spatial mental representations, supporting H1. External representation had no effect response time, providing no support for H2, and there was no interaction between external representation and distance on response time ($F < 1.5$ in both cases). Spatiality, groundedness, and physicality did not lead to faster comparisons. Contrary to the hypotheses, comparisons in the symbolic abstract condition did not require more time. Regardless of the external representation for learning energy consumption, one-by-one comparisons are relatively fast, suggesting that a floor effect that may have prevented the observation of an external representation effect. Combined with the observation of a ceiling effect on accuracy, this suggests that this type of task may have been too easy and prevented revealing all effects. Two-by-two comparisons, which were more difficult, are discussed below.

6.5.2.2 Two-by-two comparisons

In two-by-two comparisons, a 3 (distance) x 4 (external representation) repeated-measures MANOVA on accuracy and response time was conducted. The same planned repeated contrasts were used to test the effect of external representation (spatiality, groundedness, and physicality). The MANOVA revealed a main effect of distance on accuracy, $F(2, 154) = 122$, $p < .001$, $\eta_p^2 = .613$, indicating that, as found in one-by-one comparisons (Figure 11), greater distances led to higher accuracy (Figure 13), supporting H1. No main effect of external representation on accuracy was found, $F(3, 77) = 1.81$, $p = .153$, $\eta_p^2 = .066$ and no interaction between external representation and distance on accuracy ($F < 1.5$ in both cases). Compared to one-by-one comparisons, accuracy dropped at close distances (3 and 6 units) in two-by-two comparisons. The higher difficulty of two-by-two comparisons probably allowed revealing this effect. Further, although accuracy seems to be higher in the symbolic abstract condition (Figure 13), no significant results were found: The contrast corresponding to spatiality was not significant.

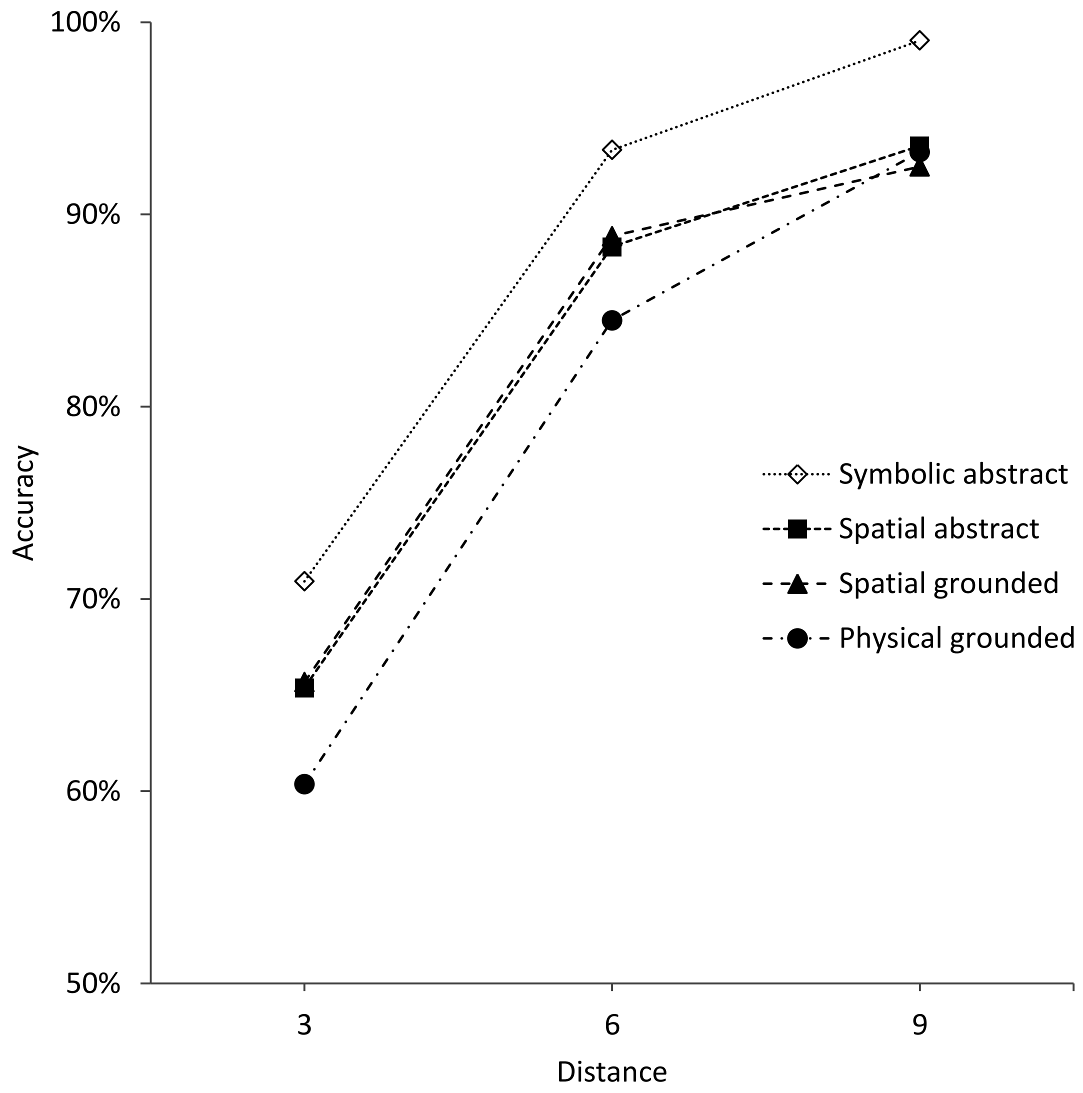


Figure 13. Accuracy for two-by-two comparisons as a function of distance and external representation.

The analysis also revealed a distance effect on response time, $F(2, 154) = 20.4$, $p < .001$, $\eta_p^2 = .209$, indicating that, as found for one-by-one comparisons (Figures 11 and 12), greater distances led to faster comparisons (Figure 14) with higher accuracy (Figure 13) and providing support for H1. This shows that comparisons were conducted with spatial mental representations. This time, the MANOVA also revealed a main effect of external representation on response time, $F(3, 77) = 4.79$, $p = .004$, $\eta_p^2 = .157$. The main effect of external representation on response time was further analyzed in repeated contrasts. The only significant individual contrast of external representation was between the symbolic abstract and the spatial abstract conditions ($p = .002$), revealing that a symbolic external representation let to longer response time than a spatial external representation, thus supporting H2b but not H2c and H2d (Figure 14).

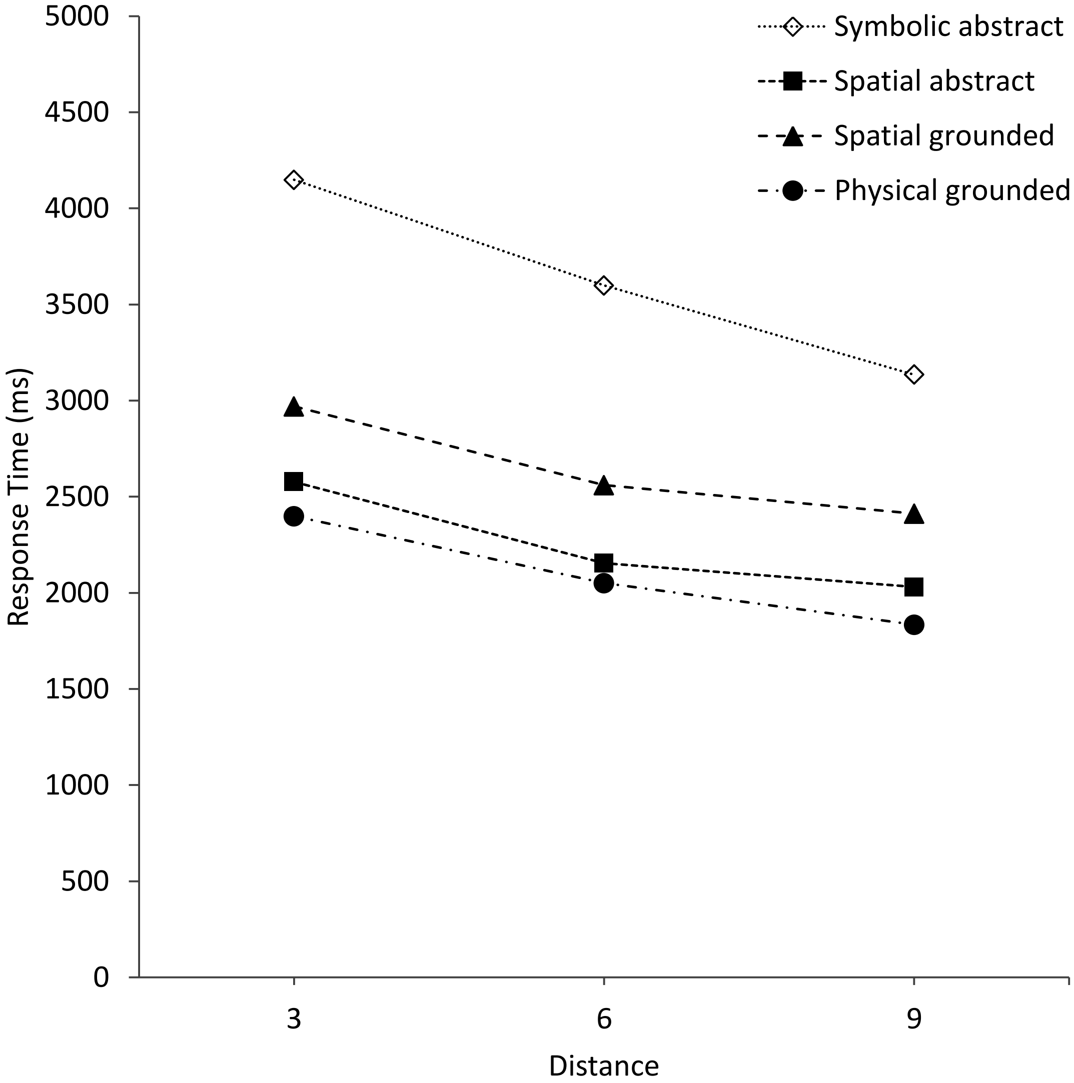


Figure 14. Response time for two-by-two comparisons as a function of distance and external representation.

This shows that learning with a symbolic external representation, as compared to spatial ones, allows similarly accurate comparisons but at the cost of speed. This can be explained by the mental generation of a spatial mental representation for comparisons from the representation stored in memory. In the symbolic condition, this requires a conversion of the mental representation from symbolic to spatial and thus at least one execution of a retrieve-convert-add-compare cycle (see Chapter 5). There was no interaction between external representation and distance on response time ($F < 1.5$). Therefore, a single cycle, rendering approximate conversions, seemed to be sufficient for two-by-two comparisons regardless of distance. Despite an apparent interaction, the contrast for spatiality and distance was not significant.

6.6 Discussion

The present study shows that external representations used in learning magnitudes of energy consumption affect performance at recall and comparison. Results suggest that the nature of an external representation, particularly whether it is symbolic or spatial, is decisive regarding the ability to later perform tasks about this magnitude. A magnitude learned with a symbolic external representation led to higher performance in recall and lower performance in mental comparison as compared to the same magnitude learned with any other spatial representation. Groundedness and physicality did not affect either recall or comparison. Implications for education and design are discussed in this final section.

6.6.1 Recalling magnitudes

The present results show that recall was more accurate when magnitudes were learned with a symbolic rather than spatial external representation, as hypothesized (H2a). With the symbolic external representation, recall was virtually perfect across the experimental group, whereas with the physical grounded external representation, the group with lowest recall performance, recall was in average sixteen percent off. Different causes may lead to this result. First, the symbolic abstract external representation provides perfect precision, and many participants indeed performed all recalls in this condition without a single error. The spatial representations, however, are intrinsically imprecise, featuring "scalar variability" or "noise" (Gallistel & Gelman, 2000). Another possible cause of the higher recall performance enabled by the symbolic abstract external representation is that the association of symbols together, a silhouette and a numeral or pair thereof, may be easier than the association of a symbol to a magnitude. When learning the magnitudes of energy consumption in either spatial condition, participants only saw variations of magnitude, whereas numeral, in the symbolic condition, differed in their shape and may thus be easier to discriminate and associate to the corresponding silhouette. This would be a surface-level processing of symbols, ignoring magnitude. However, participants across all conditions were equally able to perform the first comparison task, suggesting that magnitude was in fact processed and learned in the symbolic condition as well. Results suggest that symbolic representations are more adapted when precise knowledge and recall is required.

6.6.2 Comparing magnitudes

Theories of numerical cognition defend that magnitude comparisons are conducted with spatial mental representations. The present results support this view. The first observation in favor of a spatial mental representation of magnitude came from the distance effect observed by Moyer and Landauer (1967) . This distance effect was observed in this study, consistently with hypotheses (H1). The distance effect was equally present in all conditions of external representation, meaning that magnitude was mentally processed spatially in all cases, regardless of the external representation used in learning.

Another effect related to the spatial processing of magnitude is the Spatial Numerical Association of Response Codes (SNARC) effect, which is characterized by faster responses to larger stimuli with the right hand and conversely. This effect is attributed to the horizontal orientation of a mental magnitude representation, the number line, with larger magnitudes on the right. In this study, the SNARC effect was not observed. According to the mental representation of magnitude used, the SNARC effect can be reversed or observed vertically; in sum this effect is only present when the mental representation of magnitude causes it. There is, however, no reason why the present experimental design would lead to an absence of SNARC effect. The external representations used were either numerals, which have been shown to induce a SNARC effect, or horizontal spatial external representations with larger magnitudes on the right. This is consistent with the orientation of the number line. It is likely that this study did not have the sensitivity required to detect it, with excessive variance due to long response times. Mental comparisons were conducted from memory, contrarily to studies observing the SNARC effect where tasks were conducted in the presence of the stimuli.

Finally, spatial external representations led to faster but equally accurate mental comparisons as compared to the symbolic external representation. This is consistent with the idea that mental magnitude comparisons are conducted with spatial mental representations. It also reproduces previous results showing that external representations used in learning magnitudes affect magnitudes comparisons conducted later from memory, and generate persistent representations in memory which share their properties (chapter 5). Results also suggest that, although the spatial representations used were not continuous but discrete, their spatial arrangement in a line led to them being processed as a spatial continuum.

Grounding the spatial external representation with context and realistic perceptual cues (drawings of lightbulbs instead of dots) did not affect accuracy or response time at mental comparisons; H2c is not supported. Given the mixed evidence on this topic, it appears that grounding, in the way it was conducted in the present study, does not affect mental representation and mental processing of magnitude nearly as strongly as spatiality does.

Similarly, the physical grounded external representation led to the same results as the spatial grounded external representation regarding accuracy and response time at the mental comparisons (H2d not supported). This suggests that the importance of groundedness or physicality in representations of magnitude is negligible as compared to the importance of spatiality.

6.6.3 Practical implications

Three implications can be drawn from this study. First, symbolic rather than spatial external representations should be used to convey numerical information destined for precision and perfect recall rather than estimation, comparison, and similar magnitude processing. Second, learning with spatial external representations supports magnitude processing regardless of the external representation being realistic, contextualized, graphical, or physical. Importantly, a spatial external representation is not only a representation placed on a spatial object like a paper or screen, it is also a representation which holds spatial consistency across time and space. For instance, a line graph that would erratically scale up and down when new data is entered, or a bar graph that would not start at zero, do not constitute spatial representations; they are, however, graphical. Third, efforts to ground an external representation or to provide multiple modalities of interaction such as touch or sound would not always prove worthwhile. Emphasis should be put on spatial external representations.

6.6.4 Limitations

One limitation of this study is that it could not be ensured that all participants learned magnitudes equally well while providing the same treatment to all. As a consequence, some participants did not acquire the necessary knowledge of the appliances' energy consumption, which showed in their ability to conduct the recall and comparison tasks. Such participants were removed from the sample, as indicated in the method section. The number of participants differed across conditions, with four participants being removed from the physical grounded condition, as compared to, for instance, only one participant in the symbolic abstract condition. This study's results can thus only be generalized to learners who succeed at acquiring sufficient knowledge of magnitudes.

Another limitation is the impossibility to determine the cause of error in both recall and comparison tasks. One cause of error could be that the magnitude was recalled too high or too low, that is, an error in precision. Another cause of error, however, could be that the magnitude was properly recalled but attributed to the wrong silhouette. This would constitute an error in accuracy, a mismatch. The role of precision and accuracy and the effect of external representation on them could be further explored.

6.6.5 Future research

The present study was conducted with six appliances as materials, all of which had an energy consumption between 4 and 19 units. Future research should determine whether the present results hold for more complex and ecological situations, with values of energy consumption covering three orders of magnitude or more, as do appliances in real life—from LED lightbulbs to clothes dryers. Furthermore, other dimensions of representation than those approached in this study could be considered. The role of the body in the present study was minor, and embodiment, in association with spatiality, could reveal important findings, notwithstanding present results.

7

GENERAL DISCUSSION

In this thesis, it was first considered how energy is taught and learned about in school in terms of the concept itself and of its units. Two incompatible definitions of energy can be identified. In the scientific definition, energy is a mathematical abstraction that requires the identification of a frame of reference. In this definition, energy is always conserved according to the law of conservation of energy and always gets degraded because of entropy according to the second law of thermodynamics. In the societal definition, energy is a quasi-material substance that can be created and destroyed. In this definition, contrary to the scientific one, energy can be used up. Misconceptions arise from the simultaneous existence of the two definitions. Additionally, energy units also are a source of confusion, because of their inappropriate scale for daily use and because of the physical dimensions their names improperly evoke.

Second, it was observed that home energy management systems (HEMSs) are not designed with human cognition in mind. In order to achieve behavior change towards energy-saving behaviors, for instance using social comparison or goal-setting, external representations of energy need to be understood by energy users. Notably, problems of graphical consistency have been identified, preventing the understanding of the spatial dimension or other dimension as directly related to an amount of energy. External representations are often mere graphically enhanced tables, always requiring users to look at the axes, and taking too little advantage of spatiality in the representation.

In order to address the question of adequate external representations for dealing with energy consumption in everyday situations, human cognition was then considered. In the grounded cognition models followed in this thesis, human cognition of numerical quantities is based on mental simulations of perception and action. This led to the hypotheses that in order to learn and compare energy consumption, mental simulations of perceptual objects could be used. The literature pointed towards the importance of spatiality, and the possible benefits of groundedness and physicality. Two empirical studies were conducted to address these hypotheses, revealing original findings of multiple implications and applications.

7.1 Findings

Together, the two studies revealed findings pertaining to the mental processing of magnitude, namely learning, recalling, and comparing, with regard to external representation and comparison distance.

7.1.1 Learning can be done with both symbolic and spatial external representations

In graphs and charts, magnitude is often represented only graphically. In learning materials, however, such isolated spatial external representations never seem to be used nor to have been studied. Symbolic external representations based on numerals are advantageous learning materials because numerals are discrete and allow precision and because they are associated with verbal representations which allow communication. A learner knowing only the graphical representation of a magnitude may be only able to communicate it graphically, never verbally. Notwithstanding this limitation, spatial external representations of magnitude nonetheless allow magnitudes to be learned and mentally processed in the total absence of any symbolic representation, even indirect. In the two studies, data were screened to identify participants who failed to learn the materials, who were a minority in all conditions (see Limitations, below, for more details). The analog external representation proposed in study 1 was a rectangular vertical blue bar, which did not feature a numerical scale on the side. Thus, even indirectly, no symbolic external representation was presented, and learning was successful. The analog external representation proposed in study 2 was a horizontal array of yellow colored circles. The two studies presented in this thesis show that magnitude can be learned from spatial external representations.

7.1.2 Learning as measured by recall is more accurate with symbolic representations than with spatial representations

Empirically studying the learning of magnitudes with symbolic and spatial external representations seems never to have been done before. Recall results show in both studies that learning with symbolic representation allows for more accurate recall. There are three possible explanations for this result. First, symbols could be more easily told apart and associated with a silhouette than variations on a spatial dimension. Indeed, bars of different lengths do not present as obviously distinctive features as different Arabic numerals do.

Consequently, a part of recall error could be due to the attribution of the precise and correct energy consumption of an appliance to the wrong silhouette, which could happen less often in the symbolic conditions. A second explanation could be that, because symbolic is discrete, higher accuracy at recall may in fact be due to higher precision. Third, the recall modality in both studies was identical to the learning modality. Consequently, recalling with symbolic external representations also allows perfect precision, which is not the case when recalling with analog external representations. In the materials of the first study, a pilot was conducted to determine how much imprecision could be expected from the analog external representation, which was found to correspond to a large part of the observed error, but not all of it. The remaining difference had to be from learning and not recalling. The two present studies were not designed to distinguish between accuracy and precision, thus both may contribute to recall error.

7.1.3 All mental comparisons are conducted with analog mental representations

The distance effect in tasks of magnitude comparison is characterized by shorter response time for larger distances between the compared magnitudes, and is an indication of analog processing (Dehaene, 1992; Moyer, 1973). In the two studies of this thesis, the distance effect was observed on both accuracy and response time, regardless of external representation, and in both simple and complex tasks. It constitutes evidence of analog processing of mental comparisons in all these cases. Also, it is an indication of the quality of the experimental design. Observing this expected effect shows that the studies were sensitive enough to detect some effects on both accuracy and response time.

The distance effect is even observed on accuracy in simple comparisons in study 2, although a ceiling effect seems to take place. This is because within a set of elements to be compared, there are always simpler comparisons. For instance, comparisons involving endpoints, the smallest and largest magnitude of the set, can benefit from categorization processes such as those described by Kosslyn and colleagues (1977). However, similar to the “restriction of range” problem in correlational research, if only a certain set of distances is considered, then the distance effect can disappear within that set.

7.1.4 Distance also affects accuracy

The distance effect is generally observed on response time (Moyer & Bayer, 1976), and seems never to have been reported on accuracy. However, in both studies, it was observed on both response time and accuracy. The probable reason for this observation is that magnitude comparison was in the present studies conducted from memory, whereas the original paradigm uses direct comparisons of stimuli. In this case, accuracy would not even be measured, because error is virtually impossible. This finding suggests that distance could affect other variables, such as confidence in comparisons for instance, which could be observed in future research given the right experimental design.

7.1.5 External representations affect memorial representations

The literature is unclear regarding memorial representations, i.e. the form of mental representations stored in long term memory. Studies on imagery definitely show that visual imagery can be used as a memorial representation of objects (Denis, 2008; Moyer, 1973). However, regarding magnitude, much less certainty was found. The theory of mental simulations (Barsalou, 2008; M. Wilson, 2002) suggests that the format of memorial representations could favorably follow the format of external representations. Both present studies explicitly used mental comparisons conducted from memory to investigate the format of memorial representations given external representation. Results showed that external representations can transfer their format to memorial representations, with symbolic external representations apparently inducing symbolic memorial representations and spatial external representations apparently inducing spatial memorial representations. Given the lack of results regarding physicality in study 2, it remains unclear whether the memorial representation of a physical external representation would be a full mental simulation of this physical external representation, or if some properties, such as spatiality, would be extracted and individually kept as memorial representation.

7.1.6 Mental comparisons are slower when magnitudes are learned with symbolic representations

The difference in speed of mental comparisons of magnitude given different external representations never seems to have been studied. Models of magnitude processing indicate that mental comparisons of magnitude are conducted with analog mental representations (Dehaene, 1992). Mental representations that would not be analog would require a conversion in an analog format and thus require more time. This is indeed what is observed in the results of both studies. In the first study, in both types of comparisons, response time is significantly lower in the symbolic condition. In the second study, this is only observed in the more complex type of comparisons (a floor effect may have taken place in the simple comparisons), showing that symbolic external representations led to slower comparisons.

In the second study, the spatial representations were actually made out of a discrete number of elements (dots or pictures of lightbulbs). However, faster responses in these conditions, as compared to the symbolic condition, suggest that, as expected, the spatial external representations were perceived as analog rather than discrete. Indeed, numerosity translates to magnitude on a continuum via an accumulator (chapter 4). It must be noted, however, that instructions specified to not count the dots.

Comparisons also tended to be less accurate when learning was conducted with a spatial external representation as compared to a symbolic external representation. This was actually only significantly observed in complex comparisons in study 1, so it may depend on the case. This lower accuracy with analog external representations may simply be due to a lack of accuracy in the memorial representation, corresponding to the lower accuracy at recall. This could be compensated by deeper learning of the analog external representations, however all results may be different with deeper learning. This could be explored in future research. It is also possible that analog external representations always lead to faster but less accurate comparisons. Analog external representations would thus be more adapted for tasks that do not always need perfect accuracy but are preferred to be cognitively lighter, that is, not requiring a lot of cognitive time dedicated on the task and operations such as conversion of representations. The differences in response time between symbolic and analog are indeed quite large in complex comparisons. Saving so much cognitive time could well be worth a little more error, if it happens to occur.

7.2 Theoretical implications

The present findings confirm Dehaene's (1992) view that magnitude comparisons are conducted with analog mental representations. Indeed, a distance effect was observed in both studies, in both simple and complex comparisons, and on both accuracy and response time. The distance effect is the signature of analog mental processing (Moyer & Bayer, 1976). The findings further indicate that the distance effect can be observed on accuracy when magnitude comparisons are conducted from memory, which seems never to have been documented before.

The findings are also consistent with the mental simulations hypothesis stemming from grounded cognition, according to which the basis of all human cognition lays in perceptual and motor mental simulations (Barsalou, 2008; M. Wilson, 2002). In this view, mental representations of magnitude are perceptual simulations, consisting of mental perceptual activations directly based on the memory of external representations, or resulting from the generation of new forms of representation. For instance, after learning a magnitude with numerals as external representation, one could directly mentally simulate these numerals in order to recall them on paper for instance. However, this mental simulation would not enable mental comparisons because mental comparisons are conducted with perceptual processes and thus require mental representations with analog properties. Thus, using the magnitude in mental comparison would first require mentally simulating the numerals, then converting the representation by generating a corresponding analog perceptual mental simulation. If, on the contrary, the magnitude had been learned with an analog external representation, simply mentally simulating it would allow mental comparison without conversion and thus sparing time. This constitutes a satisfying explanation of the results of both studies reported in this thesis.

Observing faster comparisons in experimental conditions of grounded and physical external representations would have provided support for the mental simulations hypothesis, however this was not the case. Grounded and physical representations did not have any significant effect. Given that significant results were observed in all analyses, the absence of effect of groundedness and physicality cannot simply be attributed to a lack of statistical power. The absence of effect of groundedness suggests that mentally simulating a spatial arrangement of abstract objects is equally difficult as mentally simulating a similar arrangement of familiar objects, and that both can constitute equally valid mental representations of magnitude. The absence of effect of groundedness also suggests that context does not greatly ease the generation of visual-spatial mental simulations. Finally, the absence of effect of physicality suggests that different modalities of mental representations cannot convergence to form an integrated multi-modal mental representation. As in the race model of cognition according to which only one, i.e. the fastest, of either imagery or categorical processes solves mental comparisons, depending on the speed of each (Kosslyn et al., 1977), it is possible that only a single modality, the most cognitively appropriate, is used as mental representation of magnitude. Given that results in the physical experimental condition were on par with results in the spatial conditions, spatial mental representations would be, following this logic, the favored mental representations for magnitude comparisons.

Following the mental simulation hypothesis further opens a large array of questions for future research. First, according to the triple-code model (Dehaene, 1992), numbers are only represented in three different representations, one of them being a number line. However, in the framework of mental simulations, virtually an infinity of magnitude representations could exist, including for instance embodied representations (Moeller et al., 2012). It is already clear that the triple-code model is incomplete. For instance, it does not account for the multiple verbal representations, or number words, that an individual could hold in multiple languages, or similarly, different numerical alphabets. The triple-code model also describes that the analog mental representation is a number line, but as argued in chapter 4, other analog representations exist. In the first study presented in this thesis, for instance, there is no reason to believe that a number line was used by participants who learned magnitudes with a vertical bar. Their faster responses are indeed attributed to the fact that they did not need to convert their mental representations in order to conduct comparisons: a simulation of the original external representation was sufficient, thus faster. These participants arguably did not use a horizontal number line, rather vertical bars akin to the external representation with which they learned. In consequence, the mental simulation hypothesis is not only parsimonious, it also provides a more complete picture of the processes involved in mental comparisons and the variety of mental representations that could exist. This conclusion could be explored in future research, notably by seeking patterns of response facilitation in accordance with presented external representations (and thus, supposed mental representations). The Spatial-Numerical Association of Response Codes (SNARC; Dehaene et al., 1993) effect is one example of such pattern, with responses involving larger items being conducted faster with the right hand, because on the number line larger numbers are on the right. The SNARC effect can be observed vertically or inverted according to experimental settings (D. Bächtold et al., 1998; Rusconi et al., 2006). With other mental representations, other patterns could be observed.

Finally, the use of perceptual and motor simulations in magnitude comparisons would also mean that bodily states and actions, exploiting the same processes, could lead to interferences observable in response time patterns. The size congruity paradigm (Henik & Tzelgov, 1982) presented in chapter 4 and the classic Stroop task (Stroop, 1935) are examples of paradigms revealing such interferences. Furthermore, effects of bodily movement have already been observed on spatial-numerical associations (U. Fischer et al., 2016). Observing interferences could further show the reliance of magnitude processing on motor and perceptual processes.

7.3 Tension between laboratory studies and actual field

The materials and settings in the two studies represent the middle ground between controlled laboratory studies and messy real life settings. Obstacles were faced which may inform the design of future research seeking to combine experimental control and ecological design.

7.3.1 Operationalization differences

In the first study, response time in simple comparisons was affected by both distance and external representation, and by their interaction. Conversely in the second study, response time in simple comparisons was only affected by distance. The comparison task being almost identical across studies, and no difference between participant samples being expected, the absence of an effect of external representation and of an interaction is conjectured to be due to differences in the way symbolic and analog external representations were operationalized. For instance, the symbolic external representation in study 1 consisted of three-digit numbers, but of single-digit and two-digit numbers in study 2. Similarly, the analog external representation in study 1 consisted of rectangular vertical blue bars and the conceptually equivalent spatial abstract external representation in study 2 consisted of horizontal arrays of yellow colored circles. This highlights that categories such as "symbolic" and "analog" are not homogenous. Within categories, differences can exist that may lead to differences in observable effects. In other words, each of the two studies contained its own N = 1 sample of an operationalization of a symbolic external representation, among many possibilities of symbolic external representations. External representations in study 2 actually featured a superior level of descriptive detail, considering groundedness and physicality in all external representations. However, many other variables could have been taken into account, such as the number of digits, the distances used in comparisons, the numerical alphabet—and these are only examples for symbolic external representations. Future research on external representations, feedback design, and multimedia education should thus seek to operationalize categories in multiple different ways, in order to enable generalizable conclusions. A single operationalization of a category cannot be assumed to encompass the entire category.

7.3.2 Floor and ceiling effects

A floor effect on response time seems to have taken place in the simple comparisons of study 2, accompanied by a ceiling effect on accuracy. Because it was not observed in study 1, the interpretation is that, in study 2, simple comparisons were too easy. Materials contained only six appliances, as compared to eight in study 1, making them more distinguishable from one another, and magnitudes of energy consumption differed by a regular step. Conversion from symbolic memorial representation to analog transient representation was also probably easier in the second study because smaller numbers were used, in a range where people might be more fluent at manipulating them and more familiar with the magnitude they represent. Moreover, group size was smaller in the second study, lowering statistical power and preventing the observation of significant effects and rendering the data more sensitive to floor and ceiling effects. This highlights that generic laboratory tasks can be too easy, which can hide effects and lead to false negative findings. Here, if only the simple type of comparisons had been considered, the important results of this study would not have been observed.

7.4 Limitations

A first limitation of the findings of these studies is that they only apply to successful learners. In both reported studies, participants were filtered according to a threshold of both accuracy at recall and accuracy at comparisons. Participants who did not reach the threshold on either variable were considered to have failed to learn the materials and were not considered in the analyses. Future studies could further explore whether such learners present different patterns of performance, and what variables affect learning difficulty. In study 2, for instance, the filter removed a different number of participants across conditions, with a minimum of one and a maximum of four. Although these are small numbers, they beg the question of the properties of experimental materials and procedure that may influence learning.

Also, although the efforts made to present an applied situation to participants, ecological issues must be noted in the two studies and might limit the scope of the findings. First, only up to eight appliances were used per study, with values ranging within one order of magnitude. In reality, many more appliances are used in typical homes, and they range in energy consumption within three orders of magnitude or more. A modern LED lightbulb, for instance, uses about 10 watts, whereas a clothes dryer uses about 4000 watts. Some energy users may also want to consider the stand-by mode energy consumption of their appliances, which often revolves around ½ to 2 watts. In consequence, the landscape of energy consumption is for the average energy user much more complex than what was presented in the two reported studies. More complexity may affect people's ability to learn and compare magnitudes of energy consumption. Second, the appliances used in the studies were fictional in order to allow systematic manipulation of magnitudes and distance comparisons. Real appliances could not be used because they would have implied variability in prior knowledge across participants regarding energy consumption. Future applied research could investigate the extent and variability of this knowledge and its effect on energy management.

7.5 Implications for energy management

A first implication of this research is that using graphical representations is a good idea if they have enough consistency to be manipulated and are not just graphical but have relevant properties. Grounded and physical is possible a waste of time.

Also, because it is possible to learn with graphical representations, it means that whenever graphical representations are presented, learning may happen, be it intended by the designer or not. Consequently, graphical representations should always be exemplar, and display the spatial consistency they need in order to be used later for mental tasks. This can be explained with the example of the cyclist given in chapter 3: When a cyclist looks at the topographic profile of a route before or after riding, she can learn what a slope on the profile means for the slope on the road, but only if topographic profiles are spatially consistent with one another. The same slopes on the profile must indicate the same slopes in reality.

Finally, the confusion surrounding units of energy, highlighted in chapter 2, invites HEMSs designers to be careful when using these units, given the poor understanding that lay citizens have of them. Three alternative options are possible. First, it is possible to use no unit at all, or only a graphical unit. For instance, a certain amount of energy could correspond to a certain length or area of a graphical object. This is what was done in the first study of this thesis, chapter 5. However, this does not allow for easy communication between individuals. A second option would be to use a unit based on a metaphor, in any sort of fuel equivalent, like amount of coal. This would be particularly appropriate for energy users whose mix of energy sources actually contain a lot of such tangible fuels. Finally, a third option would be to use the unit proposed in chapter 2, maybe with the creative and catchy name that it lacks so far.

7.6 Implications for teaching and learning about energy

As discussed in chapter 2, different routes can be taken to educate the citizen of tomorrow and help them articulate the scientific and societal definitions of energy. However, given the importance of energy in society, it may be important to prioritize the societal definition of energy. This would be achieved by using the metaphor that energy is as quasi-material substance. Analog external representations of energy may support this definition, because they evoke the metaphor that energy is a quasi-material substance, with analog properties such as "more is up" observed when making a pile of objects. Using such external representations may thus encourage thinking of energy in the material metaphor, and could be useful in teaching the societal definition of energy.

Regarding the ability for learners to understand amounts of energy, the findings of both studies show that memorial representations, that is, knowledge as it stays in the mind, are directly built from what is learned, and that the format of representation is transferred. Learning with analog external representations would thus support all mental processes relying on such representations: estimation and comparison (Dehaene, 1992). As described in this thesis, comparison is central in the idea of making sense of a magnitude, because they only exist as compared to other ones. Estimation is however also very important, and this also the case in the scientific definition. Proper estimations allow the identification of errors in calculations, estimating the veracity of the result of computations.

Finally, the model of human cognition based on mental simulations (Barsalou, 2008; M. Wilson, 2002) was of central importance in the elaboration of the present work. Beyond specific results regarding particular representations or learning situations, this thesis shows the relevance of this model to education.

8

REFERENCES

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In *Action control* (pp. 11–39). Springer.

Andres, M., Seron, X., & Olivier, E. (2007). Contribution of hand motor circuits to counting. *Journal of Cognitive Neuroscience*, *19*(4), 563–576.

Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage. *Journal of Law, Economics, and Organization*, *29*(5), 992–1022.

Bächtold, D., Baumüller, M., & Brugger, P. (1998). Stimulus-response compatibility in representational space. *Neuropsychologia*, *36*(8), 731–735.

Bächtold, M. (2014). L'équation Elibérée=| Δm| c^2 dans le programme et les manuels de Première S. *RDST. Recherches en didactique des sciences et des technologies*, (10), 93–121.

Bächtold, M., & Munier, V. (2014). Enseigner le concept d'énergie en physique et éduquer à l'énergie : rupture ou continuité ? (pp. 21–29). Presented at the Huitièmes rencontres scientifiques de l'ARDiST, Marseille.

Bächtold, M., Munier, V., Guedj, M., Lerouge, A., & Ranquet, A. (2014). Quelle progression dans l'enseignement de l'énergie de l'école au lycée? Une analyse des programmes et des manuels. *RDST. Recherches en didactique des sciences et des technologies*, (10), 63–91.

Backlund, S., Gyllenswärd, M., Gustafsson, A., Ilstedt Hjelm, S., Mazé, R., & Redström, J. (2007). Static! The aesthetics of energy in everyday things. In *Proceedings of Design Research Society Wonderground International Conference 2006*.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89.

Ball, D. L. (1992). Magical hopes: Manipulatives and the reform of math education. *American Educator: The Professional Journal of the American Federation of Teachers*, *16*(2).

Baranes, R., Perry, M., & Stigler, J. W. (1989). Activation of real-world knowledge in the solution of word problems. *Cognition and Instruction*, *6*(4), 287–318.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660.

Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *358*(1435), 1177–1187.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(1), 617–645.

Bétrancourt, M., Ainsworth, S., de Vries, E., Boucheix, J.-M., & Lowe, R. K. (2012). Graphicacy: Do readers of science textbooks need it? In E. de Vries & K. Scheiter (Eds.), *Staging knowledge and experiences: How to take advantage of representational technologies in education and training?* (pp. 37–39). Grenoble, France.

Braithwaite, D. W., & Goldstone, R. L. (2013). Integrating formal and grounded representations in combinatorics learning. *Journal of Educational Psychology*, *105*(3), 666–682.

Brannon, E. M. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology*, *16*(2), 222–229.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1831–1840.

Cheng, P. C.-H. (1999). Unlocking conceptual learning in mathematics and science with effective representational systems. *Computers & Education*, *33*(2), 109–130.

Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *The Journal of the Learning Sciences*, *14*(2), 161–199.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. *International Handbook of Research on Conceptual Change*, 61–82.

Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994). From things to processes: a theory of conceptual change for learning science concepts. *Learning and Instruction*, *4*, 27–43.

Cohen, A. S. (n.d.). *Brainscape's "confidence-based repetition" methodology*. Retrieved from https://www.brainscape.com/images/cms/research/Confidence-Based_Repetition.pdf

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: abstract or not abstract? *Behavioral and Brain Sciences*, *32*(3–4), 313–328.

Colonnese, D., Heron, P., Michelini, M., Santi, L., & Stefanel, A. (2012). A vertical pathway for teaching and learning the concept of energy. *Review of Science, Mathematics and ICT Education*, *6*(1), 21–50.

Cornelissen, G., Pandelaere, M., & Warlop, L. (2006). Cueing common ecological behaviors to increase environmental attitudes. In *Persuasive Technology* (pp. 39–44). Springer.

Cousineau, D., & Chartier, S. (2010). Outliers detection and treatment: a review. *International Journal of Psychological Research*, *3*(1), 58–67.

Darby, S. (2006). The effectiveness of feedback on energy consumption. *A Review for DEFRA of the Literature on Metering, Billing and Direct Displays*, *486*, 2006.

de Hevia, M. D., Girelli, L., & Macchi Cassia, V. (2012). Minds without language represent number through space: origins of the mental number line. *Frontiers in Psychology*, *3*.

de Hevia, M. D., Vallar, G., & Girelli, L. (2008). Visualizing numbers in the mind's eye: The role of visuo-spatial processes in numerical abilities. *Neuroscience & Biobehavioral Reviews*, *32*(8), 1361–1372.

de Vries, E. (2001). Les logiciels d'apprentissage : panoplie ou éventail ? *Revue Française de Pédagogie*, *137*, 105–116.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1–2), 1–42.

Dehaene, S. (1997). *The number sense: how the mind creates mathematics*. New York: Oxford University Press.

Dehaene, S. (2001). Précis of the number sense. *Mind & Language*, *16*(1), 16–36.

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends in Cognitive Sciences*, *7*(4), 145–147.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371.

Denis, M. (2008). Assessing the symbolic distance effect in mental images constructed from verbal descriptions: A study of individual differences in the mental comparison of distances. *Acta Psychologica*, *127*(1), 197–210.

Dijkstra, K., & Post, L. (2015). Mechanisms of embodiment. *Frontiers in Psychology*, *6*.

DiSalvo, C., Sengers, P., & Brynjarsdóttir, H. (2010). Mapping the landscape of sustainable HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1975–1984). ACM.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2–3), 105–225.

Domahs, F., Moeller, K., Huber, S., Willmes, K., & Nuerk, H.-C. (2010). Embodied numerosity: implicit hand-based representations influence symbolic number processing across cultures. *Cognition*, *116*(2), 251–266.

Doménech, J. L., Gil-Pérez, D., Gras-Martí, A., Guisasola, J., Martínez-Torregrosa, J., Salinas, J., … Vilches, A. (2007). Teaching of energy issues: A debate proposal for a global reorientation. *Science & Education*, *16*(1), 43–64.

Duit, R. (1984). Learning the energy concept in school –empirical results from the Philippines and West Germany. *Physics Education*, *19*(2), 59–66.

Duit, R. (1987). Should energy be illustrated as something quasi-material? *International Journal of Science Education*, *9*(2), 139–145.

Egan, C. (1999). Graphical Displays and Comparative Energy Information: What Do People Understand and Prefer? In *Proceedings, European Council for an Energy-Efficient Economy*.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976a). *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976b). *Manual for kit of factor referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Elder, R. S., & Krishna, A. (2012). The "Visual Depiction Effect" in Advertising: Facilitating Embodied Mental Simulation through Product Orientation. *Journal of Consumer Research*, *38*(6), 988–1003. https://doi.org/10.1086/661531

Fails, J. A., Druin, A., Guha, M. L., Chipman, G., Simms, S., & Churaman, W. (2005). Child's play: a comparison of desktop and physical interactive environments. In *Proceedings of the 2005 conference on Interaction design and children* (pp. 48–55). ACM.

Faruqui, A., Sergici, S., & Sharif, A. (2010). The impact of informational feedback on energy consumption—A survey of the experimental evidence. *Energy*, *35*(4), 1598–1608.

Fechner, G. (1860). *Elements of psychophysics*. Holt, Rinehart & Winston.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314.

Feynman, R. P., Leighton, R. B., & Sands, M. (1965). The Feynman lectures on physics; vol. i. *American Journal of Physics*, *33*(9), 750–752.

Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency*, *1*(1), 79–104.

Fischer, M. H. (2008). Finger counting habits modulate spatial-numerical associations. *Cortex*, *44*(4), 386–392.

Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cognitive Processing*, *13*(1), 161–164.

Fischer, M. H., & Brugger, P. (2011). When digits help digits: spatial–numerical associations point to finger counting as prime example of embodied cognition. *Frontiers in Psychology*, *2*.

Fischer, U., Moeller, K., Class, F., Huber, S., Cress, U., & Nuerk, H.-C. (2016). Dancing With the SNARC: Measuring Spatial-Numerical Associations on a Digital Dance Mat. *Canadian Journal of Experimental Psychology*, *70*(4), 306–315.

Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

Froehlich, J. (2009). Promoting energy efficient behaviors in the home through feedback: The role of human-computer interaction. In *Proc. HCIC Workshop* (Vol. 9).

Froehlich, J., Findlater, L., & Landay, J. (2010). The design of eco-feedback technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1999–2008). ACM.

Froehlich, J., Findlater, L., Ostergren, M., Ramanathan, S., Peterson, J., Wragg, I., … others. (2012). The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2367–2376). ACM.

Galilée, M., Schwartz, N. H., & de Vries, E. (in preparation). Effect of external representation in learning and comparing numerical magnitude of physical properties.

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, *4*(2), 59–65.

Gamberini, L., Spagnolli, A., Corradi, N., Jacucci, G., Tusa, G., Mikkola, T., … Hoggan, E. (2012). Tailoring feedback to users' actions in a persuasive game for household electricity conservation. In *Persuasive Technology. Design for Health and Safety* (pp. 100–111). Springer.

Gimbert, F., Gentaz, E., Camos, V., & Mazens, K. (2016). Children's Approximate Number System in Haptic Modality. *Perception*, *45*(1–2), 44–55.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences*, *14*(1), 69–110.

Gustafsson, A., & Gyllenswärd, M. (2005). The power-aware cord: energy awareness through ambient information display. In *CHI'05 extended abstracts on Human factors in computing systems* (pp. 1423–1426). ACM.

Gyllenswärd, M., Gustafsson, A., & Bang, M. (2006). Visualizing energy consumption of radiators. In *Persuasive Technology* (pp. 167–170). Springer.

Ham, J., & Midden, C. (2010). Ambient persuasive technology needs little cognitive effort: the differential effects of cognitive load on lighting feedback versus factual feedback. In *Persuasive Technology* (pp. 132–142). Springer.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*(4), 389–395.

Henmon, V. A. C. (1906). *The time of perception as a measure of differences in sensations*. Columbia University, New York.

Hervé, N., Venturini, P., & Albe, V. (2014). La construction du concept d'énergie en cours de physique : analyse d'une pratique ordinaire d'enseignement. *RDST*, (10), 123–151.

Hinrichs, J. V., & Novick, L. R. (1982). Memory for numbers: nominal vs. magnitude information. *Memory & Cognition*, *10*(5), 479–486.

Holmes, T. G. (2009). *Eco-visualization: combining art and technology to reduce energy consumption*. ACM.

Hubbard, E. M., Diester, I., Cantlon, J. F., Ansari, D., Opstal, F. v., & Troiani, V. (2008). The Evolution of Numerical Cognition: From Number Neurons to Linguistic Quantifiers. *Journal of Neuroscience*, *28*(46), 11819–11824.

Hurst, M., & Cordes, S. (2017). Working Memory Strategies During Rational Number Magnitude Processing. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000169

IJsselsteijn, W., de Kort, Y., Midden, C., Eggen, B., & van den Hoven, E. (2006). Persuasive technology for human well-being: setting the scene. In *Persuasive technology* (pp. 1–5). Springer.

Jamieson, D. G., & Petrusic, W. M. (1975). Relational judgments with remembered stimuli. *Perception & Psychophysics*, *18*(6), 373–378.

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, *33*(11), 1419–1432.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, *14*(1), 103–109.

Joseph, J. H., & Dwyer, F. M. (1984). The effects of prior knowledge, presentation mode, and visual realism on student achievement. *The Journal of Experimental Education*, *52*(2), 110–121.

Karlin, B. (2011). Tracking and learning: exploring dual functions of residential energy feedback. In *Proceedings of the 6th International Conference on Persuasive Technology: Persuasive Technology and Design: Enhancing Sustainability and Health* (p. 10). ACM.

Kerst, S. M., & Howard, J. H. (1978). Memory psychophysics for visual area and length. *Memory & Cognition*, *6*(3), 327–335.

Kirsh, D. (2010). Thinking with external representations. *AI & SOCIETY*, *25*(4), 441–454.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*, *44*(1), 183–203.

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences*, *13*(2), 129–164.

Korvorst, M., & Damian, M. F. (2008). The differential influence of decades and units on multidigit number comparison. *The Quarterly Journal of Experimental Psychology*, *61*(8), 1250–1264.

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception & Psychophysics*, *14*(1), 90–94.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 47.

Kosslyn, S. M., Murphy, G. L., Bemesderfer, M. E., & Feinstein, K. J. (1977). Category and continuum in mental comparisons. *Journal of Experimental Psychology: General*, *106*(4), 341.

Kosslyn, S. M., & Pomerantz, J. R. (1977). Imagery, propositions, and the form of internal representations. *Cognitive Psychology*, *9*(1), 52–76.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago and London: The University of Chicago Press.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. Basic Books.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100.

Laski, E. V., & Dulaney, A. (2015). When prior knowledge interferes, inhibitory control matters for learning: The case of numerical magnitude representations. *Journal of Educational Psychology*, *107*(4), 1035.

Lautrey, J., & Mazens, K. (2004). Is children's naive knowledge consistent? A comparison of the concepts of sound and heat. *Learning and Instruction*, *14*(4), 399–423.

Le Postillon. (2016, December). Hydrogène : désamorcer la pompe à conneries. *Le Postillon*.

Lee, H.-S., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education*, *94*(4), 665–688.

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, *40*.

Libertus, M. E., Starr, A., & Brannon, E. M. (2014). Number trumps area for 7-month-old infants. *Developmental Psychology*, *50*(1), 108.

Löfström, E., & Palm, J. (2008). Visualising household energy use in the interest of developing sustainable energy systems.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*.

Lund, K., & Bécu-Robinault, K. (2010). La reformulation multimodale et polysémiotique comme aide à la compréhension de la physique. In A. Rabatel (Ed.), *Analyse sémiotique et didactique des reformulations* (pp. 211–244). Presses universitaires de Franche-Comté.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1.

Marshall, P. (2007). Do tangible interfaces enhance learning? In *Proceedings of the 1st international conference on Tangible and embedded interaction* (pp. 163–170). ACM.

Martin, A. (2007). The representation of object concepts in the brain. *Annu. Rev. Psychol.*, *58*, 25–45.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*, *11*(2), 194–201.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, *32*(1), 1–19.

Mazens, K., & Lautrey, J. (2003). Conceptual change in physics: children's naive representations of sound. *Cognitive Development*, *18*(2), 159–176.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 320.

Midden, C., & Ham, J. (2009). Using Negative and Positive Social Feedback From a Robotic Agent to Save Energy. New York, N.Y.: ACM Press.

Millar, R. (2005). *Teaching about energy*. University of York, Department of Educational Studies York, UK.

Moeller, K., Fischer, U., Link, T., Wasner, M., Huber, S., Cress, U., & Nuerk, H.-C. (2012). Learning and development of embodied numerosity. *Cognitive Processing*, *13*(1), 271–274.

Moreno, R., Ozogul, G., & Reisslein, M. (2011). Teaching with concrete and abstract visual representations: Effects on students' problem solving, problem representations, and learning perceptions. *Journal of Educational Psychology*, *103*(1), 32–47.

Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, *13*(2), 180–184.

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, *8*(2), 228–246.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520.

Muller, D. A. (2015). *Snatoms! The Magnetic Molecular Modeling Kit*. Veritasium. Retrieved from https://www.youtube.com/watch?v=He30D8M5fNc

Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, *50*(2), 162–188.

Nordine, J., Krajcik, J., & Fortus, D. (2011). Transforming energy instruction in middle school to support integrated understanding and future learning. *Science Education*, *95*(4), 670–699.

Núñez, R. (2006). Do real numbers really move? Language, thought, and gesture: The embodied cognitive foundations of mathematics. In *18 unconventional essays on the nature of mathematics* (pp. 160–181). Springer.

Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart, & Winston.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition*, *3*(6), 635–647.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 259–303). Hillsdale, NJ: Lawrence Erlbaum Associates.

Peirce, C. S. S. (1906). Prolegomena to an apology for pragmaticism. *The Monist*, *16*(4), 492–546.

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2013). Computational Grounded Cognition: a new alliance between grounded cognition and computational modeling. *Frontiers in Psychology*, *3*.

Pierce, J., Fan, C., Lomas, D., Marcu, G., & Paulos, E. (2010). Some consideration on the (in) effectiveness of residential energy feedback systems. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems* (pp. 244–247). ACM.

Pierce, J., Odom, W., & Blevis, E. (2008). Energy aware dwelling: a critical survey of interaction design for eco-visualizations. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat* (pp. 1–8). ACM.

Poincaré, H. (1902). *La science et l'hypothèse*. Flammarion.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211–227.

Pye, M. (2014). *The Edge of the World: How the North Sea Made Us who We are*. Penguin UK.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*(1), 1.

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Quinn, H. R. (2014). A physicist's musings on teaching about energy. In *Teaching and learning of energy in K–12 education* (pp. 15–36). Springer.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510.

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688.

Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, *99*(2), 113–129.

Sarama, J., & Clements, D. H. (2016). Physical and Virtual Manipulatives: What Is "Concrete"? In *International Perspectives on Teaching and Learning Mathematics with Virtual Manipulatives* (pp. 71–93). Springer.

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction*, *19*(6), 481–494.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*(2), 141–156.

Shaki, S., & Fischer, M. H. (2012). Multiple spatial mappings in numerical cognition. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 804–809. https://doi.org/10.1037/a0027562

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703.

Shiraishi, M., Washio, Y., Takayama, C., Lehdonvirta, V., Kimura, H., & Nakajima, T. (2009). Using Individual, Social and Economic Persuasion Techniques to Reduce CO2 Emissions in a Family Setting. New York, N.Y.: ACM Press.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–444.

Siegler, R. S., & Lortie-Forgues, H. (2014). An Integrative Theory of Numerical Development. *Child Development*, *8*(3), 144–150.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–250.

Smeaton, A. F., & Doherty, A. R. (2013). Persuading consumers to reduce their consumption of electricity in the home. In *Persuasive Technology* (pp. 204–215). Springer.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, *3*(2), 115–163.

Solomon, J. (1983). Learning about energy: How pupils think in two domains. *European Journal of Science Education*, *5*(1), 49–59.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, *19*(1), 97–140.

Strengers, Y. (2008). Challenging comfort & cleanliness norms through interactive in-home feedback systems. In *Proceedings of Pervasive 2008 Workshop on Pervasive Persuasive Technology and Environmental Sustainability* (pp. 104–108).

Strengers, Y. (2011). Designing eco-feedback systems for everyday life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2135–2144). ACM.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Thompson, J. M., Nuerk, H.-C., Moeller, K., & Cohen Kadosh, R. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychologica*, *144*(2), 324–331.

Tiberghien, A., Cross, D., & Sensevy, G. (2014). The evolution of classroom physics knowledge in relation to certainty and uncertainty. *Journal of Research in Science Teaching*, *51*(7), 930–961.

Tiberghien, A., & Megalakaki, O. (1995). Characterization of a modelling activity for a first qualitative approach to the concept of energy. *European Journal of Psychology of Education*, *10*(4), 369–383.

Trbovich, P. L., & LeFevre, J.-A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition*, *31*(5), 738–745.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology*, *18*(1), 37–54.

Van Dam, S. S., Bakker, C. A., & Van Hal, J. D. M. (2010). Home energy monitors: impact over the medium-term. *Building Research & Information*, *38*(5), 458–469.

Vince, J., & Tiberghien, A. (2012). Enseigner l'énergie en physique à partir de la question sociale du défi énergétique. *Review of Science, Mathematics and ICT Education*, *6*(1), 89–124.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4*(1), 45–69.

Vosniadou, S. (2002). On the nature of naïve physics. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 61–76). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Vosniadou, S. (2012). Reframing the classical approach to conceptual change: Preconceptions, misconceptions and synthetic models. In *Second international handbook of science education* (pp. 119–130). Springer.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488.

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping*, *31*(10), 1459–1468.

Warren, J. W. (1983). Energy and Its Carriers: A Critical Analysis. *Physics Education*, *18*(5), 209–12.

Watts, D. M. (1983). Some alternative views of energy. *Physics Education*, *18*(5), 213.

Wever, R., van Kuijk, J., & Boks, C. (2008). User-centred design for sustainable behaviour. *International Journal of Sustainable Engineering*, *1*(1), 9–20.

Wilhite, H., Hoivik, A., & Olsen, J.-G. (1999). Advances in the use of consumption feedback information in energy billing: the experiences of a Norwegian energy utility. In *Proceedings, European Council for an Energy-Efficient Economy*.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, *4*.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review*, *8*(1), 44–57.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636.

Wood, G., & Fischer, M. H. (2008). Numbers, space, and action – From finger counting to the mental number line and beyond. *Cortex*, *44*(4), 353–358.

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science Quarterly*, *50*(4), 489.

Yun, R., Lasternas, B., Aziz, A., Loftness, V., Scupelli, P., Rowe, A., … Zhao, J. (2013). Toward the design of a dashboard to promote environmentally sustainable behavior among office workers. In *Persuasive Technology* (pp. 246–252). Springer.

Zapico, J., L., Turpeinen, M., & Brandt, N. (2009). Climate persuasive services: changing behavior towards low- carbon lifestyles. New York, N.Y.: ACM Press.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179–217.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*(1), 87–122.

Zhang, J., & Patel, V. L. (2006). Distributed cognition, representation, and affordance. *Pragmatics & Cognition*, *14*(2), 333–341.

9

APPENDICES

9.1 Appliances used in experimental studies (chapters 5 and 6)

Name: Sock-and-Roll
Used in: Bedroom
The Sock-and-Roll folds, organizes, and dispenses the inventor's socks according to the forecasted weather.

Name: Absolute Ruler
Used in: Workshop
The Absolute Ruler measures and controls the size and weight of objects and parts used in the inventor's work.

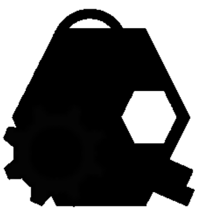
Name: Epic Bubbler
Used in: Bathroom
The Epic Bubbler generates a deluge of soap bubbles to make every shower unforgettable.

Name: Pantry Pilot
Used in: Pantry
The Pantry Pilot deftly rotates and organizes the inventor's goods so that food is never forgotten and never goes bad.

Name: Smart Spice
Used in: Kitchen
The Smart Spice automatically grinds and mixes spices according to the smell and flavor of cooking meals.

Name: Rainbow Maker
Used in: Kid's Room
The Rainbow Maker makes crayons of any color imaginable, with or without glitter.

Name: Document Den*
Used in: Office Room
The Document Den scans, stores, and organizes paper documents to make them easily available on demand.



Name: Tool Trapper*
Used in: Garage
The Tool Trapper snatches any tool forgotten on the floor when the inventor repairs his vehicles.

* not used in the study presented in chapter 6.

9.2 Comparison distances for chapter 5

This first table presents the distances between the energy consumption of appliances when compared one-by-one. In bold is the energy consumption of individual appliances, as presented with the symbolic external representation. Crossing these values gives the distance. One dot indicates a Close distance, two dots Medium, three dots Far.

	147	**184**	**230**	**287**	**359**	**449**	**561**	**701**
147	0	-37	-83	-140	-212	-302	-414	-554
184	37.	0	-46	-103	-175	-265	-377	-517
230	83.	46.	0	-57	-129	-219	-331	-471
287	140.	103.	57.	0	-72	-161	-274	-414
359	212..	175..	129.	72.	0	-90	-202	-342
449	302…	265..	219..	161..	90.	0	-112	-252
561	414…	377…	331…	274..	202..	112.	0	-140
701	554…	517…	471…	414…	342…	252..	140.	0

This second table presents the distances between the energy consumption of appliances when compared three-by-three. Not all possibilities of comparisons were used, only the present selection. On the right is indicated the distance corresponding to comparisons. Coding is the same as above.

184	**287**	**561**	×	**230**	**359**	**449**	→	6.
147	**701**	**359**	×	**184**	**449**	**561**	→	13.
230	**147**	**561**	×	**184**	**287**	**449**	→	18.
230	**287**	**561**	×	**184**	**359**	**449**	→	86.
184	**287**	**359**	×	**230**	**147**	**561**	→	108.
287	**359**	**449**	×	**184**	**230**	**561**	→	120.
184	**359**	**561**	×	**230**	**287**	**449**	→	138.
287	**147**	**561**	×	**184**	**230**	**359**	→	222..
230	**287**	**701**	×	**184**	**359**	**449**	→	226..
230	**359**	**561**	×	**184**	**287**	**449**	→	230..
184	**230**	**287**	×	**359**	**449**	**147**	→	254..
184	**287**	**561**	×	**230**	**359**	**701**	→	258..
184	**359**	**561**	×	**230**	**147**	**449**	→	278..
230	**287**	**359**	×	**184**	**449**	**561**	→	318…
184	**230**	**449**	×	**287**	**359**	**561**	→	344…
287	**359**	**701**	×	**184**	**230**	**561**	→	372…
184	**287**	**359**	×	**230**	**449**	**561**	→	410…
287	**449**	**561**	×	**184**	**230**	**359**	→	524…
184	**230**	**287**	×	**359**	**449**	**561**	→	668…
184	**230**	**449**	×	**701**	**359**	**561**	→	758…

9.3 Comparison distances for chapter 6

This first table presents the distances between the energy consumption of appliances when compared one-by-one. In bold is the energy consumption of individual appliances, as presented with the symbolic abstract external representation. Crossing these values gives the distance.

	4	**7**	**10**	**13**	**16**	**19**
4	0	-3	-6	-9	-12	-15
7	3	0	-3	-6	-9	-12
10	6	3	0	-3	-6	-9
13	9	6	3	0	-3	-6
16	12	9	6	3	0	-3
19	15	12	9	6	3	0

This second table presents the distances between the energy consumption of appliances when compared two-by-two. All possibilities of comparisons were used. On the right is indicated the distance corresponding to comparisons.

16	**13**	×	**19**	**7**	→	3
7	**19**	×	**10**	**13**	→	3
7	**10**	×	**4**	**16**	→	3
19	**4**	×	**16**	**10**	→	3
13	**10**	×	**4**	**16**	→	3
13	**4**	×	**7**	**16**	→	6
19	**7**	×	**16**	**4**	→	6
16	**10**	×	**13**	**7**	→	6
16	**10**	×	**19**	**13**	→	6
10	**19**	×	**7**	**16**	→	6
13	**7**	×	**4**	**10**	→	6
4	**19**	×	**16**	**13**	→	6
10	**4**	×	**7**	**16**	→	9
13	**7**	×	**10**	**19**	→	9
13	**19**	×	**7**	**16**	→	9