Universal Design for Learning (UDL) is a framework mainly focusing on the concept of variability in learning, providing guidelines for content creation, class teaching and assessment. Previous literature reviews suggest that there is a limited amount of evidence about its effectiveness, and in general about other frameworks based on Universal Design. So an exploratory meta-analysis was performed in order to investigate the trends in the existing amount of research evidence about the tested effectiveness of UDL.

A collection of abstracts was obtained from the ERIC database, from which N=80 peer reviewed journal articles were considered. Abstracts were classified by 5 parameters: a) the presence of explicit results in the text; b) the positiveness of results; c) the category which benefited from the results (students, teachers or both), d) the sample size; e) the country where the study took place. This review allowed to clarify some points: research evidence about the effectiveness of the UDL framework is present in a minority of the retrieved abstracts, but the amount of available information is growing; position and communication papers are more frequent than research reports, which is probably due to the need of guidelines by school personnel; research is mainly going on in the US. The analysis presented here includes a reflection on its own limitations and on how it could be continued.

KEYWORDS: Universal Design for Learning, inclusion, accessibility, variability, instructional content
Introduction

Universal Design for Learning (UDL) is a theoretical framework focusing on the concept of variability in learning. It provides guidelines for content creation, class teaching and assessment based on research evidence mainly coming from the educational neuroscience field (Dolan et al. 2005). It was originally developed at the Center for Applied Special Technology (CAST), an educational research & development organization based in Wakefield, MA. UDL provides a set of guidelines, which have been published and reviewed by CAST through time; the 1.0 version was released in 2008 and was then updated to version 2.0 in 2011. Prior to this public release, however, UDL was already known to the public since the 1980s, as can be assumed from CAST’s timeline.

UDL is based on brain science research, and specifically on the existence of three main neural networks (Rose & Meyer 2002) that are involved in the learning process:

1. recognition networks – the “what” of learning – gather facts from the world around us and categorize what we see, hear, and read;
2. strategic networks – the “how” of learning – organize and express our ideas, allowing to perform tasks related to problem solving or content production;
3. affective networks – the “why” of learning – connect the learning experience to feelings and emotional background, determining engagement and motivation.

UDL’s main goal is to help teachers and educators in being sensible to the differences characterizing every student (e.g. those with learning disabilities, or those with lower motivation, as well as the very proficient ones). In order to do so, it theorizes three principles, directly descending from the functions of the three networks mentioned above. The three principles are:

1. provide multiple means of representation, presenting information and content in different ways;
2. provide multiple means of action and expression, differentiating the ways that students can express what they know;
3. provide multiple means of engagement, stimulating interest and motivation for learning.

The UDL framework is highly connected with educational technology, as it is seen as the most convenient way to comply with the above listed principles. By the way, some authors (King-Sears 2009) stress the fact that its pedagogical component must be equally considered. However, it is widely seen as a useful means to foster the so called “technology enhanced learning”. It is seen as a useful solution for reducing learning barriers and supporting the needs of all learners, as it “adds a philosophical structure to technology-enhanced courses that can change the education landscape and create a more dynamic learning experience for all involved” (Morra & Reynolds 2010).

Each UDL principle contains three guidelines that specifically address one aspect of that domain. In turn, guidelines include up to four checkpoints, which provide
tips and guidance about different learning tasks (CAST 2008). Every checkpoint was defined on the basis of research evidence: CAST’s website provides a list of research papers that led to the definition of their recommendations. This supports the validity and the robustness of the principles and guidelines proposed by UDL, but leaves some space open for questioning the effectiveness of this method, its applicability and its sustainability. Although the neuroscientific evidence at the base of the guidelines is not discussed here, an exploratory meta-analysis is presented, quantifying the existing amount of research evidence about the tested effectiveness of UDL.

The rest of this paper is organized as follows: in the “literature review” section are presented references to the more general Universal Design paradigm, from which UDL in some ways descends. A distinction is also made between UDL and a similar framework called Universal Design for Instruction. The following section illustrates the method used in this meta-analysis, which was designed to be simple and easily repeatable. Generic results, as well as more specific findings, are discussed. The “discussion” section goes through the findings in order to clarify some interesting points and to set the basis for future research.

1. Literature review

The name and main idea of “Universal Design for Learning” is inspired to the well known, established paradigm of Universal Design (UD), coming from the field of architecture and industrial design. The expression was originally coined by architect Ronald Mace and was later referred to “products, environments, programmes and services to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design” (NDA 2009). The basic idea behind UD is that it is more convenient to create environments that are accessible from the beginning, rather than to modify and fix them with specific, later interventions (Preiser & Smith 2010). UD has derivatives in different domains: UDL and other frameworks such as Universal Design for Instruction (Scott et al. 2003) represent and recontextualise its principles and basic ideas in the field of instructional content creation. As already highlighted by Guglielman (2011), there is a difference in the way UDL and UDI address the same issue: while UDL concentrates on the three neural networks described above and translates their functions into principles, UDI reflects on the seven principles of Universal Design (Center for Universal Design, 1997) in an educational fashion and adds two more principles: “a community of learners” and “instructional climate”. Nevertheless, some authors tend to use the names of UDL and UDI as interchangeable (Koch et al. 2006).

Orr and Bachman Hammig (2009) focus on UD-related approaches in supporting students with Learning Disabilities (LD): in their study they examine 38 research papers gathered through different channels, using combinations of 9 keyword phrases. In their classification of research evidence the less represented topic is what they call backward design, or the clarity in setting goals that for Rose and

1 http://www.udlcenter.org/research/researchevidence.
Meyer (2002) is “the essential first step in teaching”. Another literature review by Roberts et al. (2011) on the topic of Universal Design for Instruction in post-secondary education contexts suggests that the amount of empirical evidence about the effects of UDI implementations is still limited. The same seems valid for UDL: the study presents data about it and in general about UD applications, with an overall low number of retrieved research articles. The amount of papers about UDL is of 18 across four databases, but this seems to include an unspecified number of duplicates.

2. Method and results

2.1 Research design

A collection of abstracts was obtained from the ERIC database, which states to provide “unlimited access to more than 1.4 million bibliographic records of journal articles and other education-related materials, with hundreds of new records added multiple times per week”. A previous study (Roberts et al. 2011) based on ERIC and on three other sources obtained more valid results from it than from the others combined. A query was submitted for the descriptor “universal design for learning” (quotation marks included). The search yielded a total of 141 records, among which only N=80 were taken into consideration as they represented peer reviewed journal articles. The search was performed at the end of year 2012 and the results discussed here refer to the articles available at the date of 8 December 2012.

The collected abstracts were read and classified by 5 parameters:

a) the explicit presence of results in the text: this is considered a simple indicator of how much structured and non-structured data is available;
b) the positiveness of such results: as a research result could confirm or confute a theory, abstracts were classified according to their stating the success of the UDL framework or not;
c) the category who benefitted of the results (students, teachers or both): in the case of the above described “positive results” it is interesting to know who benefitted of them, in order to determine whether there could be preferable fields of application or areas to be better investigated;
d) the sample size: this parameter tells how much differentiated is the range of the retrieved research results; as sample composition largely depends on methods, it reflects the variability in methodological approaches used;
e) the country where the study took place: this is considered useful in order to estimate to what extent UDL practices are known and studied outside the area where they were originally developed.

2 The other three databases used by Roberts et al. (2011) were Academic Search Premier, PsycInfo, Social Sciences Citation Index.
2.2 Sample description

The publication time for the collected sample of abstracts spans between 2000 and 2012. Those explicitly mentioning research results are 19 (23.75% of the total), while the others are mainly position papers or literature reviews, and two abstracts introduce research on best practices. Figure 1 shows the distribution of articles per year: simple visual inspection leaves no doubt about the increasing popularity of this particular topic. There is a peak in the number of abstracts and research results in 2008, the same year when the public 1.0 version of UDL guidelines was released. This event probably fostered the high article production of 2009 and 2010. A second peak in research results is visible in 2011, in correspondence with the release of the 2.0 version of the guidelines.

![Figure 1. Amount of articles per year, distinguishing those with research results from the rest.]

2.3 Findings

The following paragraphs provide insights about the five classification criteria described above.

*Explicit presence of results in the text:* the 19 abstracts mentioning research are the only ones cited in the “bibliography” section among those examined. More than half of them appeared in the 2007-2012 period, a fact that can be interpreted as an increased attention towards research evidence for the UDL framework. The presence of empirical results also increased through time, even if it seems to have slowed down in 2010.

*Positiveness of results:* successful results of UDL interventions are highlighted in all the 19 abstracts mentioning and explicitly quantifying them. Apart from them, in 12 other abstracts there is no mention of an empirical research, yet a positive evaluation is given, through sentences like “this study measured changes and/or improvements in instruction as perceived by students”, without telling the magnitude of the improvements, or what method was used for assessment. The other abstracts (n=49) provide no mention of experimental results or other observations.
The results are classified and summarized in table 1: only the 19 abstracts with explicit results are taken into account, and as some of them reported more than one type of result, the total of the figures in the “number of papers” column is higher.

<table>
<thead>
<tr>
<th>Type of result</th>
<th>Number of papers</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>methodological improvement</td>
<td>11</td>
<td>Improvement in how teacher prepare content for their classes; feeling of self-efficacy</td>
</tr>
<tr>
<td>academic improvement</td>
<td>4</td>
<td>Improvement in how students perform in assessments and standard tests.</td>
</tr>
<tr>
<td>improved knowledge of problems</td>
<td>2</td>
<td>Improvement in consciousness about what barriers exist and how much widespread problems are</td>
</tr>
<tr>
<td>students appreciation</td>
<td>2</td>
<td>Positive feedback from students about the courses they follow</td>
</tr>
<tr>
<td>digital divide decrease</td>
<td>1</td>
<td>Improvement in accessibility of websites</td>
</tr>
<tr>
<td>environmental improvements</td>
<td>1</td>
<td>Improvement in the personalization of the classroom environment</td>
</tr>
<tr>
<td>social improvement</td>
<td>1</td>
<td>Improvement in the social relationships between students</td>
</tr>
</tbody>
</table>

Table 1. Results classified by type in the 19 abstracts explicitly mentioning a measured result

Category who benefited of the results: among the 31 abstracts that report a beneficial result, the majority address students (45.16%), fewer teachers (25.81%), while some others involve both categories (22.58%).

Sample size: The size of the samples involved in the analysed studies is highly varied: it ranges from three students in one case to hundreds of them, as shown in table 2. They could be homogeneous (addressing only one kind of subjects), mixed or loosely defined.

Country where the study took place: The research described in the collected abstracts was mainly carried out in the US (n=27), with the only exceptions of 2 abstracts referring to Singapore, one to Australia and one to Brunei. As the UDL framework was created by a research center based in Massachusetts, this result is not particularly surprising. It is interesting, though, that no papers referred to research performed in Europe, where Universal Design is a well known paradigm, while its derivatives UDL and UDI are probably not so much popular.
This review allowed to clarify some points about the amount and composition of research evidence about the UDL framework. As a first result, research evidence about the effectiveness of the UDL framework is confirmed to be limited, as only a small percentage of the retrieved abstracts points to research results. Yet an important information can be added to this previous knowledge: the amount of available information is growing at an interesting pace: abstracts with results were 0.66 per year in the 2000-2005 period, while they are 2.66 per year between 2006 and 2011. The majority of the results highlight improvements in teaching practices, which are among the core objectives of UDL, together with academic results on the side of students and consciousness about problems. The release of the first version of UDL guidelines in 2008 was accompanied by a strong increase in the amount of research evidence, and the same happened in 2011. The period in the middle was characterized by a strong growth in publication.

Table 2. Study samples grouped by composition. The data comes from the 19 abstracts with explicit results and from two abstracts mentioning a sample but no results

<table>
<thead>
<tr>
<th>Homogeneous samples</th>
<th>Mixed samples</th>
<th>Generically defined samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>36 + 5 teacher candidates</td>
<td>16 students with significant intellectual disabilities in Grades K-2, 9 teachers</td>
<td>adult stakeholders at five schools nominated as having exemplary inclusive HSBLPs³</td>
</tr>
<tr>
<td>41 teachers</td>
<td>16 teachers, 1153 students</td>
<td>Participants from nine states and four countries</td>
</tr>
<tr>
<td>867 students in Grades 5-12</td>
<td>271 faculty members and teaching associates (TAs); 92 additional faculty members and TAs. 98 faculty members and administrators</td>
<td>preschool children with disabilities</td>
</tr>
<tr>
<td>five university professors</td>
<td>Focus groups (n=57) and surveys (n=665) of students with disabilities and faculty members</td>
<td>schools</td>
</tr>
<tr>
<td>representative consumer health websites</td>
<td>student teaching portfolios, formal teacher observations, and 6th- through 12th-grade student assessments</td>
<td>students with disabilities, students who are English language learners and typically developing students</td>
</tr>
<tr>
<td>75 pre-service special education teachers</td>
<td></td>
<td>students surveyed and interviewed during and after a course</td>
</tr>
<tr>
<td>ten high school students with learning disabilities</td>
<td></td>
<td>students with special needs</td>
</tr>
<tr>
<td>three students with multiple disabilities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>two groups: students with high incidence disabilities and general education students</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3 High School Service Learning Programs.
Position and communication papers, produced all along the 13-years period considered here, account for the majority of the analysed sample: this is probably due to the need of guidelines by teachers, educators and school personnel.

The performed search was designed to be rigorous and easily repeatable, but also has some limitations: first of all, it was limited to one database only. Second, full text articles were not always available. Faced with the impossibility to retrieve the full text of every paper, we decided to consider only abstracts. This may have introduced bias, mainly because of the lack of a standardized method for writing them.

Future steps in this research will include an extension of the abstracts collection using more databases, in order to prevent bias due to the possible incompleteness of the sources. It would also be useful to determine if the proportion between abstracts explicitly reporting results and generic ones is a constant: in such a case, this could be a convenient sampling technique for analysing and comparing the diffusion of different frameworks that operate on the same level.

As outlined in the “findings” section of this paper, UDL seems to be only applied in the US, with a few exceptions. As Universal Design in general is a largely known and discussed topic, it is our intention to investigate the presence of traces of related research in the rest of the world, and specifically Europe.

References


