Received: 20 April 2020 | Revised: 28 January 2021 | Accepted: 2 February 2021

DOI: 10.1111/ijcs.12662

ORIGINAL ARTICLE

International Journal of Consumer Studies WILEY

Investigating the effectiveness of simplified labels for safe use communication: The case of household detergents

Maggie Geuens[1] | Dominic Byrne[2] | Geert Boeije[3] | Virginie Peeters[4] | Bert Vandecasteele[5]

[1]Faculty of Economics and Business Administration, University of Ghent, Ghent, Belgium

[2]A.I.S.E., International Association for Soaps, Detergents and Maintenance Products, Auderghem, Belgium

[3]Boeije Consulting, Sint-Martens-Latem, Belgium

[4]InSites Consulting, Ghent, Belgium

[5]InSites Consulting, Düsseldorf, Germany

Correspondence
Dominic Byrne, A.I.S.E., International Association for Soaps, Detergents and Maintenance Products, Boulevard du Souverain 165, Brussels 1160, Belgium.
Email: dominic.byrne@aise.eu

Funding information
The research described in this paper was commissioned by A.I.S.E. who funded the work of InSites Consulting and Boeije Consulting. Prof. Geuens interpreted the data and authored the publication independently, without any form of compensation from A.I.S.E.

Abstract

This study assessed the effectiveness of safety communication on the back labels of hazardous products (with regulatory and safety information as dictated by regulatory requirements), with household detergents as a test case. The potential of simplification to increase label effectiveness was evaluated by comparing the currently used labelling approach with two simplified alternatives. The labels mainly differed in terms of the amount of information and the prominence of pictograms. The generalisability of theoretical insights on the effectiveness of pictograms in safety messages to a more real-life context was tested by (a) realistic labels containing several other information elements besides the safety information and (b) target users who are knowledgeable about the product type. One thousand eight hundred (1,800) respondents participated in an online experiment and were randomly exposed to one of the labels. The positive cognitive and behavioural effects commonly attributed to pictorials could not be confirmed, but positive affective effects did emerge. Specifically, even though participants were asked to carefully read the label, they did not spend enough time to process all the content except for the most simplified label. The results did not show meaningful differences between the three labels in terms of information recall (which was poor for all executions), hazard perceptions and behavioural intentions when confronted with an accident. In contrast to this lack of differentiation in cognitive and behavioural intention effects, we did find a clear difference in the affective measure. A majority of the respondents preferred the simplified safety labels. As such, avoiding information overload, and conveying the information in an easier way by means of more prominent use of pictograms, appeared to be appreciated by consumers of household products, while it did not negatively impact label effectiveness.

KEYWORDS

CLP regulation, hazard communication, pictograms, product safety labels, simplification

1 | INTRODUCTION

1.1 | Background

In the current marketplace, consumers are overwhelmed by information. A significant part of this concerns hazard information or warnings, ranging from warnings of being exposed to dangerous substances in the household to medical risks (Hess et al., 2011). Hazard information is useful to the extent that the user actually reads the provided information. Unfortunately, this is often not the case. A limited attention span, lack of time, already having excessive information, brand trust, product familiarity, low hazard perception and perception of low personal relevance are but a few reasons for not reading information labels (Bartels et al., 2018; Dörnyei & Gyulavári, 2016; Lee & You, 2020; Moreira et al., 2019). A recent study on chemical household products, for example, showed that only 1.8% of the participants claimed to always read the information and only 18.9% reported to do so often (Lee & You, 2020). The applicable EU regulatory framework (i.e., CLP—the EU Classification, Labelling and Packaging Regulation EC 1272/2008; the EU Biocidal Products Regulation EC 528/2012; the EU Detergents Regulation EC 648/2004) dictates that specific hazard communication be included in the label of hazardous products. In relation to prescribed hazard communication requirements, a recent study showed that consumers judged that regulatory and safety labels on detergent products are overloaded with difficult to understand information (SynapsesQuali, 2016). In addition, in the Eurobarometer study of 2010 (European Commission, 2011), only half of the respondents judged that daily used products like detergents require safety instructions and only about one-quarter said that they always read the instructions that come with everyday detergents. The European Commission (2012) recommended that the content simplification of labels should be promoted (for instance providing further guidance on omitting certain information elements and on precedence rules).

Simple and easy to understand safety communication on products is important to increase the likelihood that consumers (a) notice the safety information, (b) understand it and (c) act upon the information to ensure the safe use of the product (Laughery, 2006). Too many hazard statements, overly detailed warnings and too cluttered or too difficult information may all overload consumers leaving them unmotivated to read or try to understand the information (Bialkova et al., 2013; Rhoades et al., 1990; Rogers et al., 2000). This may especially be the case for household products such as detergents which consumers use on a regular basis, are familiar with and do not perceive as very dangerous (Laughery, 2006). Familiarity has been shown to lower the perception of risk, decreasing the likelihood that consumers pay attention to and encode warning information (Dörnyei & Gyulavári, 2016; Rogers et al., 2000). Hence, the relevance of exploring the effectiveness of alternative simplified labels.

An often-advocated way to simplify risk information is the use of pictorials rather than written text (e.g., Laughery, 2006). Pictorials include, for example, actual photographs, representative drawings and abstract symbols (i.e., pictograms). Pictorials' potential has been shown in many different contexts, but the majority of previous studies (a) investigated stimuli that only or predominantly contained warning information and (b) used non-target consumers as respondents. Back-of-pack labels of real household products such as detergents do not only contain warning information, but also product usage information, dosage information, ingredient disclosure, most of which comes in text form and in multiple languages. Warning information on real-life product labels thus does not appear in isolation but competes for attention with several other information elements. In addition, and in reality, household products such as detergents are mostly used by consumers who are very familiar with the product type. Both the complexity of real product labels and accumulated user experience may attenuate pictorials' potential. Thus, the question is whether the positive effects previously observed for the use of pictorials versus text also generalise to complex information labels resembling real-life product labels for consumers who use the product on a regular basis. Starting from a representative back of pack label of a detergent product (i.e., a back label that was equivalent, in presentation and content, to a typical marketplace back labels), we aim to close this research gap by developing two simplified versions of the current label (differing in number and size of pictorials) and compare their effectiveness in a large group of target consumers.

In what follows, we first pay attention to the literature on the superiority of pictorials versus text in terms of cognitive, affective and behavioural effects. Next, we formulate our research objective. Before explaining our results, we extensively explain our methodology (participants, description of the labels under investigation, our research design and the variables we measured). We end with a general discussion that includes a discussion of our research, the limitations that characterize our study, a few suggestions for future research and our conclusions.

1.2 | The use of pictorials versus text in warnings

Even though it is not entirely clear how to optimally communicate risk (Siegrist et al., 2008), there is a surge in research that advocates simplification through the use of visual information such as pictorials (e.g., Laughery, 2006; Okan et al., 2017, 2020). Pictorials offer not only cognitive but also affective and behavioural advantages over written text. Indeed, pictorials can communicate a lot of information at a glance and can visually represent the potential hazard, the potential consequences and what people could do to prevent the hazard (Kalsher et al., 1996). Pictorials have been shown to more easily attract attention (Houts et al., 2006; Kaufmann & Ramirez-Andreotta, 2019; Niederdeppe et al., 2019; Sathar et al., 2016); to easily communicate guidelines because of the large amount of information they can capture (Wogalter et al., 2006); to be quickly comprehended (Dowse & Ehlers, 2005; Vigoroso et al., 2020); and to be understandable for everyone (also for low-literacy people—Adams et al., 2010; and for non-native speakers—Vigoroso et al., 2020; or people with impaired vision—Handcock et al., 2004). Although warning effectiveness does not always benefit from product familiarity

(Rogers et al., 2000), several studies report that product familiarity does enhance pictorial comprehension (e.g., Caffaro et al., 2018; Chan & Ng, 2010; Lesch et al., 2011). Pictorials have also been shown to facilitate the retrieval of information (Sathar et al., 2016; Wilkinson et al., 1997) and to be better remembered than words (Dalvie et al., 2014; Laughery & Wogalter, 2014; Ta et al., 2010). Thus, in general, pictorials may facilitate the cognitive processing of the information.

Research on the fluency or perceived ease with which new information can be processed, has shown that fluency increases believability, liking, preference and confidence in one's judgements (Alter et al., 2007; Fang et al., 2007; Hawkins & Hoch, 1992; Roose et al., 2019; Unkelbach, 2007; Winkielman et al., 2003). Moreover, pictorials are also often more vivid, powerful and intense than text. This has been shown to lead to stronger persuasion effects (Kees et al., 2010; Kelley et al., 1989; Witte & Allen, 2000). As pictorials are more vivid and easier to process, it is not surprising that labels with pictorials not only engender a positive cognitive effect, but also engender an affective effect. In this sense, research has shown that labels with pictorials are preferred above labels without pictorials (Kalsher et al., 1996; SynapsesQuali, 2016). A recent study, using a participatory design involving consumers in the design of communication concerning cancer risks due to environmental exposures shows that participants recommended that graphic elements should outweigh text (Kaufmann & Ramirez-Andreotta, 2019). In addition, they induce more fear (Andrews et al., 2014; Davis & Burton, 2016; Niederdeppe et al., 2019) and increase respondents' perception of danger (Boelhouwer et al., 2013; Friedmann, 1988) (to note, this may not always be desirable for daily used products without safety concerns).

Interestingly, next to cognitive and affective effects, labels with pictorials have also been shown to exert a positive effect on behavioural compliance such as smoking cessation after exposure to pictorial warnings on cigarette packs (Andrews et al., 2014; Davis & Burton, 2016; Niederdeppe et al., 2019; Noar et al., 2016). The foregoing underscores the necessity to evaluate labels not only on cognitive outcomes, but also on affective and behavioural aspects.

It is important to mention that prior research has shown that warning designs that effectively address one particular communication objective (e.g., enhance hazard understanding), may at the same time compromise others (e.g., have people adopt safe behaviour) (Okan et al., 2018). For example, when graphically communicating the proportion of a population that is at risk, showing this as a percentage of the population leads to a good risk understanding but appears to be ill-suited to promote risk-avoidant behaviour. Vice versa, graphs that show the absolute number of people at risk are more effective at driving safe behaviour, but will score poorly in terms of conveying a good understanding of the risk (Ancker et al., 2006; Stone et al., 2015, 2017). As an example in the area of cleaning products, the hazard of a drain cleaner is fundamentally different from the hazard of a hand dishwashing liquid. Using the same hazard pictogram on both product labels may result in over- or underestimating the real hazard and lead to inadequate precautionary behaviour. The challenge thus is to construct product labels that balance cognitive, affective and behavioural consequences.

Another observation is that a lot of insights have already been gathered on (a) the comprehension of different safety pictorials and warning symbols (e.g., Caffaro et al., 2018; Dalvie et al., 2014), (b) the effects of specific information displayed in graphs (e.g., Ancker et al., 2006; Okan et al., 2020), (c) the impact of design factors (such as size, colour/contrast, signal word, graphics and format), non-design factors (such as product familiarity, location and distraction) and personal factors (such as demographics) on warning effectiveness (Laughery & Wogalter, 2014; Noar et al., 2016; Sathar et al., 2016). Yet, most of the stimuli used in previous research are simplified versions of product labels used in the marketplace. That is, only a specific pictorial or only a safety warning were tested, whereas in reality the pictorial and the warning are integrated into a complex product label. To bridge theory and practice, research on the impact of pictorials in real-life product labels in product users is a necessary and straightforward step to extend this stream of research.

1.3 | Research objective

In line with the research of Laughery (2006), we aim to investigate whether simplifying information on pack, *inter alia* by a more prominent use of pictograms, could enhance consumers' understanding and liking of the safe use information, without compromising their perception of the hazards or their resulting safe use behaviour. The current study focuses on the holistic effectiveness of the back label of a laundry detergent—a typical daily used household product that is classified as hazardous. As a benchmark, the current labelling approach, fully in line with regulations, is used. Simplification is assessed via two alternative simpler labels (that still aim for maximal practical realism).

As consumers currently tend to largely ignore the information on the back label of such products and express a general dislike of this label (SynapsesQuali, 2016), this category serves as a suitable test case whether simplification of the label might lead to improved cognitive, affective and behavioural effects.

Previous research that assessed the benefits of pictorials on label comprehension and effectiveness (as outlined above) often had simplistic versions of labels evaluated by a test population different from the target users (e.g., DeJoy, 1989). In contrast, the current study aims to test realistic label designs in a target user test population:

- Real(-istic) labels necessarily contain a lot of information due to regulatory requirements and due to the use of multiple languages (a common practice in most EU countries). The study aims to assess the holistic effectiveness of a label with more pictorials relative to a current benchmark. However, both the test labels and the benchmark are more sophisticated than the ones used in previous research. In order to provide necessary product information and comply with current regulation,

several information elements need to be included (e.g., how to use the product, precautionary guidelines, ingredients, allergic information, company logo, etc.). This means that the pictorials are surrounded by many other label elements, which all compete for the user's attention. This complexity could negatively impact pictorial effectiveness.

- Target users who are already very familiar with the product because they use the product category on a regular basis, are likely to have a preconceived opinion about the product's hazards (not very hazardous in case of a laundry detergent) and usage instructions. As such it is expected that they may pay less attention to information on the label (cf. SynapsesQuali, 2016).

These aspects—that all contribute to the realism of the tested scenarios—could attenuate the favourable cognitive, affective and behavioural effects reported in prior single-variable research. Hence, the need of this investigation.

We focus on cognitive, affective and behavioural communication outcomes because this tripartite is usually reflected in communication objectives and campaign effectiveness measurements (De Pelsmacker et al., 2021). Although traditional hierarchy-of-effects models assume consumers to move, respectively, through a cognitive, affective and finally behavioural stage, more recent research acknowledges that depending on the situation a different sequence can be followed. As such, information on all three aspects is relevant to evaluate communication effectiveness (De Pelsmacker et al., 2021). Hence, our main research questions can be formulated as follows:

RQ1: Does back label simplification with pictorials (vs. a more information-dense label) lead to cognitive benefits for a familiar, frequently used product?

RQ2: Does label simplification with pictorials (vs. a more information-dense label) engender more positive affective effects for a familiar, frequently used product?

RQ3: Does label simplification with pictorials (vs. a more information-dense label) induce more positive behavioural intentions for a familiar, frequently used product?

To evaluate cognitive benefits, the recall and comprehension of safety aspects on the label are assessed. Affective effects are captured by investigating respondents' preference for the different label types and behavioural intention is measured by investigating intended behavioural compliance in case of a concrete hazard situation. We expect positive effects of label simplification with pictorials in line with the research results discussed in Section 1.2, but the familiarity with the product and the fact that—next to the pictorials—still quite some information needs to be listed to have a 'realistic label', may attenuate such positive effects.

2 | MATERIALS AND METHODS

The study aimed to assess the holistic effect of label simplification—rather than the specific impact of individual changes (such as the use of pictograms instead of or in addition to text; different ingredient label approaches; inclusion vs. omission of warnings or information). Hence, no single-variable comparisons were conducted for individual label elements—but instead, entire label concepts were evaluated—with multiple differences between the label conditions.

2.1 | Participants

The study was conducted online, with 1,812 participants, recruited from the database of InSites Consulting. Table 1 shows the demographic characteristics of the respondents.

To ensure a heterogeneous sample that reflects consumers across the EU we included equal shares respondents from four different countries (France, Poland, Spain, Sweden). This takes into account that these may have a different understanding, or may attach a different meaning to pictograms and safety labels (Klaschka & Rother, 2013). Note that the current research aimed to broadly cover the EU overall, but not to assess potential differences between countries.

TABLE 1 Demographic characteristics of the sample

	France (%)	Poland (%)	Spain (%)	Sweden (%)
	(*n* = 453)	(*n* = 451)	(*n* = 456)	(*n* = 452)
Gender				
Male	30.0	30.0	30.0	30.0
Female	70.0	70.0	70.0	70.0
Age				
18–24	13.4	15.4	11.3	13.5
25–34	19.1	23.0	23.6	18.6
35–44	21.3	18.0	23.6	20.9
45–54	20.6	20.7	19.4	18.9
55–70	25.7	23.1	22.2	28.3
Children < 18 years				
Yes	35.6	42.4	39.3	27.1
No	64.4	57.6	60.7	72.9
Household's Net Monthly Income				
Less than 600€	3.6	9.5	2.2	2.7
600€–1,499€	21.4	54.1	26.2	13.7
1,500€–1,999€	24.3	26.8	33.9	20.5
2,000€–2,999€	21.7	5.4	23.4	21.1
3,000€–3,999€	16.0	1.8	8.2	18.6
4,000€–5,000€	10.0	0.8	4.0	14.8
More than 5,000€	3.1	1.5	2.0	8.5

Only people who confirmed to be (mainly or jointly) responsible for the purchase of laundry detergents and for doing the laundry were eligible to participate. In line with this objective, a quota of 30% males/70% females was implemented.

2.2 | Label design

The 'Current label' tested in the study aimed to be representative of typical liquid laundry detergents that can be found in the EU market. Like many of these products, the hypothetical detergent in the study is classified as hazardous under the CLP Regulation for Serious Eye Irritation (Cat.2). Further, it contains one allergen at a level requiring labelling with the hazard phrase EUH208 '*Contains Alpha-isomethyl ionone, May produce an allergic reaction*'. In addition it contains other allergens that require inclusion as per Detergent Regulation ingredient reporting requirements. The classification for Serious Eye Irritation requires the labelling of the GHS hazard phrase H319 '*Causes serious eye irritation*'. and the pictogram GHS07 (exclamation mark) with the signal word '*Warning*'. It triggers the following precautionary phrases: P102 '*Keep out of reach of children*'. P305/351/338 '*IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing*'. P337/313 '*If eye irritation persists: Get medical advice/attention*'. P301/312 '*IF SWALLOWED: Call a POISON CENTRE/doctor if you feel unwell*'. and P101 '*If medical advice is needed, have product container or label at hand*'. In addition to the CLP labelling elements, additional compulsory elements are the list of ingredients and the dosage instructions, as required by the Detergents Regulation. Finally, other elements such as the logo for voluntary initiatives (A.I.S.E. charter and safe use icons) and the bar code were also added, to fully reflect the label content in reality.

To be representative of reality, the labels were multilingual. For each country in the study, a label for a relevant country cluster (the test country together with three neighbouring countries) was used, in line with common practice. This way the respondents would not be confused by seeing languages they were not used to having on product labels. In each of these country cluster executions, the test country's language was listed as the first language. The different clusters were: Poland (Polish, Czech, Slovakian, German); Sweden (Swedish, Danish, Finnish, Norwegian); France (French, Italian, Spanish, Portuguese); Spain (Spanish, Portuguese, French, Italian). The translations in the different languages were taken from the legal text where prescribed, or as necessary, were provided by experts of the relevant national detergent associations.

Next to the 'Current label', two alternative back label designs were developed (Figure 1). The design strategies are outlined in Table 2. The content of the labels was developed by an ad hoc working group consisting of detergent industry experts in regulatory toxicology and consumer communication/marketing (six individuals, including representatives of the market-leading companies). This group aligned on what is representative of the current situation (for the Current label), as well as on what information is of the highest consumer relevance and what safety guidance is the most essential, for the two Alternative labels.

2.3 | Experimental design

Online, the respondents were asked to take a close look at the front and the back label, as if they would buy this product for the very first time. They were reminded to take into account their personal situation (e.g., their own preferences, concerns, family situation). To ensure good visibility, respondents could digitally zoom in or out as they wished.

A three-group between-subjects design was set up in which participants were randomly assigned to one of the three test label conditions. In each condition, they were first exposed to a front label of a laundry detergent (identical in all three conditions), followed by a back label (i.e., the safety label). This back label was either the Current label, the Alternative 1, or the Alternative 2 label. Thus, the assessment of label information recall and comprehension was limited to one condition for each participant, to avoid biased responses driven by carry-over between the labels.

Front label (common across the three conditions)	Current label	Alternative 1 label	Alternative 2 label

FIGURE 1 Test labels [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Test label design strategies

Label elements	Current label (benchmark)	Alternative 1 label	Alternative 2 label
Overall design principle	Based on currently marketed products, meeting all regulatory requirements	Simplified, more graphical alternative, nevertheless still conveying all current messages irrespective of their relevance	Highly simplified alternative, focusing only on the most consumer relevant messages, independent of the current regulatory framework
CLP/ detergent regulation	Fully compliant	Flexible interpretation of the regulation. CLP pictogram (exclamation mark) and hazard statement are maintained	Label information not in line with current regulatory requirements. Except for the allergen phrase EUH208, no CLP text or pictograms were used
Safe use instructions	All relevant precautionary phrases (in all languages). A.I.S.E. safe use icons for "Keep away from children" and "Eye hazard" (due to space limitations, these safe use icons were relatively small).	Where possible, safe use icons replaced the corresponding precautionary text ("Keep away from children", "Eye hazard", "Seek medical help in case of eye exposure" (newly developed for the purpose of this experiment), and "Ingestion hazard") (larger than on the "Current label" option)	Two prominently sized icons to convey the safety advice judged most essential by safety experts ("Keep away from children" and "Eye hazard"). Absence of CLP text or visuals enabled sizing the retained icons about 4x larger than on the "Current label" execution
Ingredient information	Ingredients list in all languages. Allergens were listed in two separate places, as per current labels	International ingredient list (INCI, as used for cosmetics)—including all allergens—replacing the multilingual ingredient list from the detergent regulation	List with the names of allergens, next to a prominent visual cue saying "ALLERG" replacing the ingredient list

In addition, following the comprehension assessment and behavioural intention measurement, a subgroup of participants was shown all three conditions (within-subjects) and indicated their label preference (see below).

2.4 | Metrics

To assess how the three labels score on different communication objectives, we measured respondents' cognitive (reading, recall and hazard perceptions), behavioural (intentions after reading a scenario of a concrete incidence) and affective responses (label preference). Also, the time spent reading the label was tracked.

2.4.1 | Label reading time

During the online study, the time that respondents spent viewing the back label, prior to engaging with the questions, was automatically tracked. The online participants were asked to read the label to the same extent as they would do in reality, when they would buy this specific product for the first time.

SynapsesQuali (2016) reports that consumers tend to be disinterested in the back label of daily used household products such as detergents. The European Commission (2011) had made similar observations. Hence, it was expected that insufficient time might be spent to adequately study all of the label's content. To provide a benchmark for the amount of time truly required to effectively read and view all the information on the back labels, a follow-up study was conducted amongst 14 people (A.I.S.E. employees, based in Belgium). Each person was exposed to the three labels in a random sequence and was asked to read the full text (in one language) at a normal pace, in a way that the content is well understood and to also carefully look at all graphical information. The labels were multilingual (four languages), always with either the employee's native language or a fluent second language included. The labels were provided by email as JPG files, to be viewed on the screen with the possibility to zoom in—that is, similar to the viewing experience in the main experiment. The participants used a stopwatch to determine the time needed to read each label. It should be noted that reading the label to a sufficient extent is based on self-reporting. Furthermore, the test population is very familiar on a professional level with detergent products, which may also lead to a shorter reading time relative to the general public. Consequently, the reading times reported by the A.I.S.E. employees are expected to represent the low end of the true required reading time.

2.4.2 | Information recall

The 'stickiness' of the information was assessed, that is, the information that participants spontaneously remembered after viewing the label. Without the label in front of them, they described in their own words what they had seen or read on the back label. After the completion of the online research, interpreters (fluent in the language of the respondents) coded the open-ended input. For each respondent, they judged whether the answer was equivalent to one (or more) pre-defined potential answers that had been developed in advance, based on brainstorming about what one might expect

respondents to recall from the label. In case no suitable pre-defined response was available, a direct translation in English was made. This open-ended approach was used instead of presenting the respondents with a picklist of the pre-defined answers, to ensure that what they reported was not biased by seeing the options list.

2.4.3 | Hazard perceptions

Hazard perceptions were measured on 5-point Likert scales. Participants indicated how dangerous they thought the product was for their health, for children and for the environment (1 = not dangerous at all, 5 = very dangerous). More specific hazards were measured by asking to what extent the participants agreed (1 = completely disagree; 5 = completely agree) with several statements: 'You should avoid contact of this product with your eyes'. 'You should avoid contact of this product with your skin'. 'You might die if you drink this product'. 'The product may cause an allergic reaction'. And 'This product should be stored out of children's reach'. It should be noted that the statement about potential lethal toxicity was included to probe the perception of the respondents. In reality, this type of product is not classified for acute toxicity. Participants who had indicated that one of their family members suffered from allergies, asthma or sensitive skin, were also asked to what extent they agreed that the information about allergies was clearly visible on the label and easy to read.

2.4.4 | Safety intentions

As a proxy for behavioural compliance with the label's recommendations, the following situation was presented to the respondents: 'Please imagine this laundry detergent accidentally splashed into the eyes of someone in your household. This person says it's quite painful'. Without being shown the label again and without referring to the label, the respondents were asked what they would do, via an open-ended question (free text). After the completion of the questionnaire, the responses were coded by interpreters and either matched with pre-defined answers or translated in case of unique answers, in the same way as outlined above in Section 2.4.2.

Eye exposure was selected as the case study, because in practice eye splashes tend to be the most frequent type of accident (as aligned by the ad hoc industry expert working group involved in this research) and because eye irritation is the only hazard for which the product category in the study commonly has a regulatory classification.

2.4.5 | Label preference

As a subsequent step, 500 randomly selected participants from the sample (evenly divided across the four countries), were presented the three labels in parallel and were asked which one they preferred. They were also asked to indicate why they preferred this label by ticking the box(es) of the applicable reason(s). Eight possible reasons were provided to select from (as listed in Table 5), and in addition, the respondents were allowed to provide additional reasons via free text input. To note, this assessment was limited to a subset of the total test population, because the other subjects were involved in different follow-up questions outside of the scope of the current paper.

3 | RESULTS

3.1 | Label reading time

For the three different label conditions, the time that was used by the respondents in the main experiment to view the back label is provided in Table 3, next to the time required by A.I.S.E. employees to read and view all the information on these labels.

For the respondents in the main experiment, 56 outliers (with reading time > Q3 + 3xIQR) were removed (cf. Hoaglin et al., 1986). No differences emerged across label conditions in the time spent reading the label (one way ANOVA: $F(2, 1{,}753) = 2.725$, $p = .066$; all p's of the post hoc comparisons >.10) and the average time was only 29.0 s ($SD = 23.4$).

In contrast, the A.I.S.E. employees (who were aware that they had to read/view every label element in order to assess the time needed to process all this information) did show differences in the time spent reading the label (repeated measures general linear model: $F(2, 12) = 19.367$, $p < .001$). They required significantly more time for the more complex Current and Alternative 1 labels compared to the highly graphical Alternative 2 label (respectively $F(1, 13) = 41.907$, $p < .001$ and $F(1,13) = 11.357$, $p = .005$). They did not need significantly more time for the Current label (with more text) than the Alternative 1 label (with more pictograms) ($F(1, 13) = 3.652$, $p = .078$).

On average, A.I.S.E. employees needed 50.4 s ($SD = 29.3$) per label, significantly longer than the time used by the respondents in the main experiment ($t(42{,}259) = -4.693$, $p < .001$). For the individual label conditions, the reading time taken in the main experiment was significantly shorter than the time required by the A.I.S.E. employees to process all information for the Current label ($t(588) = -4.742$,

TABLE 3 Back Label reading time (seconds)

	Current Label	Alternative 1 label	Alternative 2 label
Online respondents in the main experiment	$M = 30.9$ $SD = 26.4$ ($n = 576$)	$M = 28.1$ $SD = 21.8$ ($n = 583$)	$M = 28.1$ $SD = 21.6$ ($n = 597$)
A.I.S.E. Employees	$M = 64.8$ $SD = 29.4$ ($n = 14$)	$M = 54.1$ $SD = 32.0$ ($n = 14$)	$M = 32.2$ $SD = 15.0$ ($n = 14$)

Abbreviations: M, mean, SD, standard deviation.

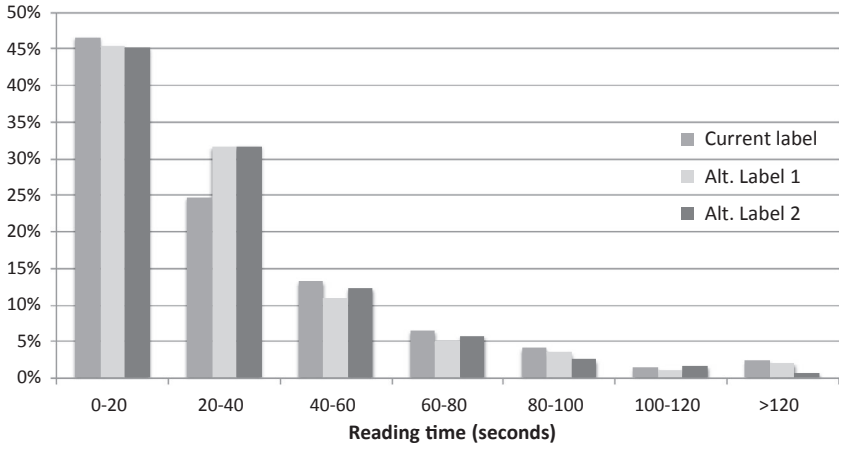

FIGURE 2 Distribution of back label reading time by the online respondents

$p < .001$) and the Alternative 1 label ($t(13{,}292) = -3.025$, $p = .010$). Moreover, for Alternative 2 label reading time did not differ between the two groups of respondents ($t(609) = -.708$, $p = .479$).

The distribution of the reading times (Figure 2) shows that nearly half of the respondents in the main experiment (approximately 45%) used less than 20 s, irrespective of the label condition. These findings suggest that the majority of the online respondents used insufficient time to study the two more complex back labels in order to process and internalize all the information present.

3.2 | Information recall

As for the information stickiness (i.e., label information that participants spontaneously remembered and reported afterwards), the most often recalled pieces of information were dosage instructions (27.8% of respondents) and washing instructions (18.6% of respondents). According to a recent Finnish study, the most used course of action to determine the detergent dosage is to follow the dosage instructions mentioned on the package (Miilunpalo & Räisänen, 2019). As the package is thus often consulted for this information (certainly the first time they use the brand), it is not surprising that this is also the information that sticks best. The recall of information elements most related to the topic of this manuscript is listed in Table 4.

A first observation was that the proportion of respondents mentioning at least one safety-related aspect did not differ between the three different labels. Differences did emerge for specific individual safety instructions (see Table 4 for the statistics of Chi-Square tests), but none of the labels showed an advantage across all elements.

The message 'Keep away from children' was recalled by twice as many respondents in the Alternative 2 label condition (prominent icon only) as in the Current label or Alternative 1 label condition (a smaller icon, respectively, with and without accompanying text). In contrast, the eyes hazard warning was mentioned by significantly more respondents for the Current label (smaller icon with hazard and precautionary text) than for the Alternative 1 label (smaller icon with hazard text only) or Alternative 2 label (prominent icon only).

TABLE 4 Percentage of respondents spontaneously recalling information elements from the labels (open-ended question, chi-square tests)

	Label			
Recalled elements	Current	Alt. 1	Alt. 2	$\chi^2(2)$
Safety and safe use instructions	16.7[a]	16.7[a]	11.5[b]	8.400*
Information about product hazard	6.3[a]	7.2[a]	3.3[b]	9.210*
Keep away from children	9.3[a]	9.7[a]	19.2[b]	33.693***
Painful/dangerous for the eyes	20.0[a]	12.7[b]	11.0[b]	22.121***
Warning: may cause allergic reactions	3.0[a]	12.0[b]	9.8[b]	34.873***
Ingredient list/product composition	10.2[a]	7.0[b]	6.0[b]	7.968*
At least one aspect related to product safety was recalled	40.3[a]	44.5[a]	43.7[a]	2.382

Note: Cell entries refer to the percentage of respondents who listed the information element. Cell entries with different superscripts refer to significant differences between the labels whereas the same superscript refers to a non-significant difference.

Abbreviations: Alt. 1, alternative label 1, Alt. 2, alternative label 2.

*$p < .05$; ***$p < .001$.

The presence of safe use instructions and information about product hazards/danger, as generic concepts, was more often reported for the Current label and Alternative 1 label than for the Alternative 2 label. However, the allergy warning was recalled by significantly more people in the Alternative 1 and Alternative 2 label conditions compared to the Current label condition.

The presence of information about the product composition was noticed slightly more for the Current label (10.2% of respondents) than for the Alternative 1 label (7% of respondents) and Alternative 2 label (6% of respondents).

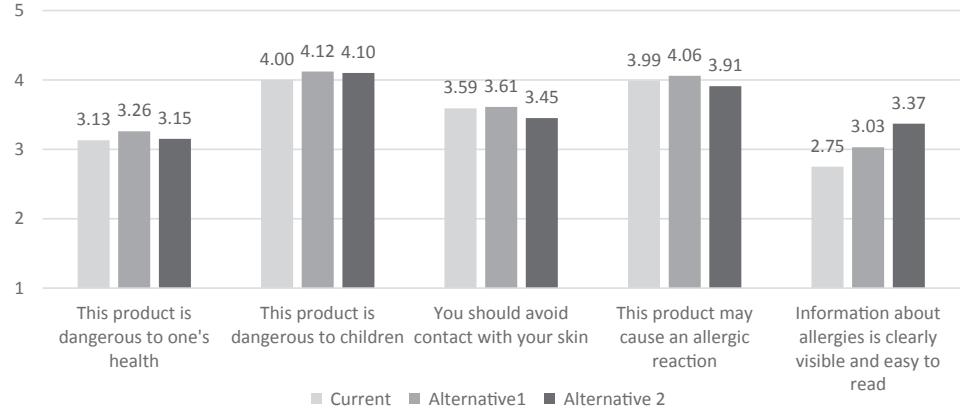

FIGURE 3 Hazard perceptions induced by the three labels (n = 600 per label execution). Dependent variables measured on a 5-point Likert scale (1 = completely disagree; 5 = completely agree). "Information about allergies is clearly visible and easy to read" was only evaluated by those respondents who had, at the start of the study, mentioned a concern with allergies, asthma or sensitive skin (n = 268, 239, 252 for respectively the Current, Alt. 1 and Alt. 2 labels)

Overall, there was no single label element that the participants systematically remembered very well. This observation was similar for the three label conditions. This is also reflected in the fact that no differences emerged in the proportion of respondents mentioning at least one safety-related aspect. Prior research (cf. above, Section 1.2) indicated that more prominent use of pictorials may increase label comprehension. In the current study, the pictorials are surrounded by text—and may have attracted less attention than in previous studies.

3.3 | Hazard perception

The hazard perceptions the respondents reported in the different label conditions were mostly very similar, except for the clarity of the allergy information (Figure 3).

Comparing the perceived danger for health between the different labels (F(2, 1,797) = 2.694, p = .068, η^2 = .003), all means were close to the midpoint of the scale, reflecting that respondents felt neutral about the hazard, rather than having the feeling that the product was dangerous. Thus, the only observed statistically significant difference—the Alternative 1 condition experienced as slightly more hazardous (M_{Alt1} = 3.26, SD = 1.03) than the Current label (M_{Curr} = 3.13, SD = 1.07; t(1,797) = 2.132, p = .033)—is in the practice of limited relevance.

Regarding the perception of how dangerous the product is for children (F(2, 1,797) = 2.506, p = .082, η^2 = .003), the means are about one point higher on the 5-point scale than the general hazard perception. Also for the children hazard, the perception was somewhat higher in the Alternative 1 condition (M_{Alt1} = 4.12, SD = .98) than in the Current condition (M_{Curr} = 4.00, SD = 1.00; t(1,797) = 2.120, p = .034). But despite this observation, the hazard rating can be considered largely equivalent across the labels.

The perception of having to avoid contact of this product with one's skin was also slightly different between the labels (F(2, 1,797) = 3.252, p = .039, η^2 = .003). It was lower when having read the Alternative 2 label (M_{Alt2} = 3.45, SD = 1.17) compared to the Current label (M_{Curr} = 3.59, SD = 1.15; t(1,797) = 2.016, p = .044) and the Alternative 1 label (M_{Alt1} = 3.61, SD = 1.17; t(1,797) = 2.361, p = .018). As with the general hazard perception, as the differences are small and the ratings were close to the neutral mid-point (means around 3.5 on a 5-point scale), practical relevance is limited.

The Alternative 2 label led to a lower perception of an allergic reaction risk than the Alternative 1 label (F(2, 1,797) = 3.218, p = .040, η^2 = .004; M_{Alt2} = 3.91, SD = 1.13; M_{Alt1} = 4.06, SD = 1.02; t(1,797) = 2.516, p = .012). Nevertheless, like for the other labels, it still resulted in a mean close to 4 on a 5-point scale ('agree').

The clarity of the allergen information, as experienced by respondents who had reported that they have a concern with allergies, asthma or sensitive skin (n = 760), was the only notable difference between the three label conditions (F(2, 756) = 13.608, p < .001, η^2 = .036). This subgroup found the allergen information to be significantly more clearly visible and easier to read on the Alternative 2 label (M_{Alt2} = 3.37, SD = 1.32) than on the Alternative 1 label (M_{Alt1} = 3.03, SD = 1.36; t(756) = 2.788, p = .005). The latter in turn scored significantly higher than the Current label (M_{Curr} = 2.75, SD = 1.41; t(1,797) = 2.308, p = .021). As the Current label scores below the midpoint of the scale, this indicates that respondents felt the allergy information is *NOT* clearly visible and easy to read.

The somewhat stronger overall health hazard perception for the label that contains the largest number of pictograms is in line with the findings reported by Boelhouwer et al. (2013) and Friedmann (1988). Nevertheless, the differences between the three labels were minimal and likely of little practical relevance. Taking into account the (too) limited amount of time respondents spent reading the label, hazard perceptions may have been based more on intuition and prior knowledge than on a real interpretation of the safety label. Important to note is that those participants for whom information on allergies, asthma and sensitive skin was most relevant, rated the simplified Alternative 2 label highest on visibility and ease of reading of this type of information.

3.4 | Safety intentions

Similar to the hazard perception, the open-ended question probing into what the participants would do when the laundry detergent accidentally splashed into the eyes of a household member, resulted in very similar answers across label conditions. Table 5 shows an overview of the most frequently mentioned answers, per condition, together with the statistical results of Chi-Square tests. A large majority of the participants (about 80% for each of the three labels) reported that they would rinse the eyes with water, which is, indeed, the recommended first measure to be taken. It is also the action that was reported for 96% of accidental eye exposures to detergent and cleaning products reported to Poison Control Centres across several countries in Europe (Scazzola et al., 2019). About one in five respondents would immediately go to or call the doctor. This was most mentioned for the Alternative 1 condition (24.2%) (significant difference vs. the Alternative 2 condition, $\chi^2(1) = 6.100$, $p = .008$). Whilst for the Current label condition, calling the doctor was the most mentioned response (11.8%) if pain or vision problems persisted (significantly more than in the Alternative 1 label condition, $\chi^2(1) = 7.011$, $p = .005$). No other significant differences emerged across the conditions.

It is striking that overall less than 3% of the participants mentioned that they would themselves consult the label and follow the instructions mentioned on the label. And only about 5% would show the label to the person who is helping.

3.5 | Label preference

Significantly more respondents preferred the simpler and more graphical Alternative labels than the Current label ($\chi^2(2) = 59.354$, $p < .001$) (see Figure 4). This is in line with the earlier qualitative observations (SynapsesQuali, 2016) that consumers dislike crowded labels with a lot of text. It is also in line with the recommendation to convey safety information in an attractive and easy to understand way (Laughery, 2006).

Table 6 provides an overview of the reasons participants indicated for their label preference (including statistics of the Chi-Square tests). Not surprisingly, for those who preferred the Current label, the main reason was that it provided the most information. The ones who preferred the Alternative 2 label mainly did so because it was easy to understand and read, and because they liked the graphical design and layout. Alternative 1, the alternative with multiple safety icons,

TABLE 5 Intended behaviour in case of eye exposure (chi-square tests)

	Label			
	Current	Alt. 1	Alt. 2	$\chi^2(2)$
Rinse the eyes (with water)	82.3	79.8	79.0	2.291
Immediately go to/call the doctor	21.5[a,b]	24.2[a]	18.3[b]	6.098*
Immediately go to the hospital/emergency room	12.5	13.3	13.3	0.245
Call the doctor if the pain or vision problem persists	11.8[a]	7.3[b]	9.2[a,b]	7.184*
Immediately call an ambulance /emergency services	8.0	9.2	7.7	0.980
Go to the hospital/emergency room if the pain or vision problem persists	6.0	5.3	6.7	0.946
Show the label to the person who is helping	5.3	5.7	3.3	4.201
Consult the label and follow the instructions	2.2	2.8	2.7	0.580

Note: Cell entries refer to the percentage of respondents who listed the behaviour. Cell entries with different superscripts refer to significant differences between the labels whereas the same superscript refers to a non-significant difference.

Abbreviations: Alt. 1, alternative label 1, Alt. 2, alternative label 2.

*$p<.05$.

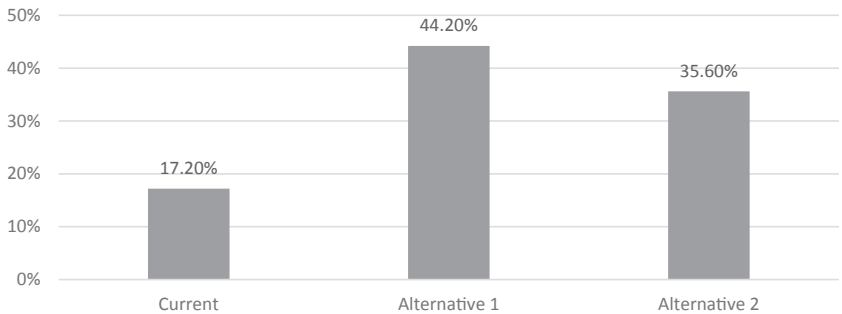

FIGURE 4 Label preference ($n = 500$)

TABLE 6 Reasons for label preference (chi-square tests)

	Label			
Reason for preference	Current	Alt. 1	Alt. 2	$\chi^2(2)$
This label provides the most information	88.4[a]	64.3[b]	32.4[c]	83.358***
The pictograms are easy to understand	8.1[a]	60.6[b]	60.1[b]	77.346***
The other labels do not contain enough information	41.9[a]	19.0[b]	6.7[c]	47.373***
The other labels contain too much information	3.5[a]	10.9[b]	22.9[c]	21.503***
The other labels are too complicated	8.1[a]	9.5[a]	24.2[b]	20.479***
This label has the nicest graphical design/layout	11.6[a]	28.1[b]	38.2[c]	20.073***
This label is easy to read	34.9[a]	54.3[b]	55.6[b]	11.401**
Information on this label is most relevant to me	27.9[a]	35.3[a]	20.7[b]	10.382**

Note:: Cell entries refer to the percentage of respondents who listed the reason for preference. Cell entries with different superscripts refer to significant differences between the labels whereas the same superscript refers to a non-significant difference.

Abbreviations: Alt. 1, alternative label 1, Alt. 2, alternative label 2.

**$p<.01$; ***$p<.001$.

bridged the gap between the two other labels. It was the label most preferred, with as main arguments that it provided a lot of information, while still being easy to understand and read.

4 | GENERAL DISCUSSION

4.1 | Discussion

Consumers judge that the current safety labels on household detergents are overloaded and that the information is difficult to understand (SynapsesQuali, 2016). The objective of the current study was to test how two simplified alternatives compare with the current label. In line with (a) Laughery (2006) who claims that product warnings should be attractive and should provide easy to understand information in order to be noticed, be understood and engender safe use behaviour—and with (b) research that advocates the use of pictorials to simplify information (e.g., Okan et al., 2018), we constructed the alternative labels in such a way that they mainly differed in how much textual information they contained and how prominent the safety pictograms were.

Our study (a) used realistic labels which are more complex than previously investigated labels as they contain several other information elements besides the warning message, (b) tested three labels (a current, representative label and two simplified alternative versions) that differ in the amount of text and the prominence of the pictograms, (c) focused on consumers who regularly use the product and are well acquainted with it and (d) included an affective and behavioural intention measure next to comprehension. As such, this study provides not only a holistic assessment on the effectiveness of simplified and more graphical safety labels, it also provides the first test of whether pictorials' proclaimed positive effects can be generalized to more real-life situations. More specifically, the question was whether also in this context, the more prominent use of safety pictograms and overall simplification would lead to a better noticing and comprehension of the safety messages (RQ1 cognitive effects), a higher preference for these simplified labels with pictorials (RQ2 affective effect) and ultimately to better compliance in terms of safety intensions (RQ3 behavioural intention effect).

Prior research warns that too many hazard statements, overly detailed warnings and too cluttered or too difficult information may leave consumers unmotivated to read (Rhoades et al., 1990; Rogers et al., 2000). And for daily used products like detergents, only a quarter of the population claims to always read the instructions (European Commission, 2011). In the current study, this was reflected by the fact that participants—even when explicitly asked to do so—did not spend enough time to read everything for the more information-rich labels (i.e., they spent less time than technically needed to be able to read and view all content). Irrespective of the safety label execution, none of the information elements were particularly well-recalled. In general, no meaningful differences were detected between the labels (neither in terms of information stickiness nor in terms of hazard perceptions). For example, the number of people recalling at least one safety element was the same (and always below 50%) for all three labels. One exception is the clarity and visibility of allergen information as reported by respondents with an allergy concern—this was significantly better for the most simplified Alternative 2 Label. The Alternative 2 Label lists the names of allergens, next to a prominent visual cue saying 'ALLERG' which replaced the full ingredient list. Respondents with an allergy concern appreciated this visual cue together with only the relevant information much more than the full information depicted in the other two labels. This is not surprising as the more information appears on a label, the more clutter is created and the less attention-getting properties a single information element gets (Bialkova et al., 2013). Ineffective allergen communication causing confusion due to its complexity and ambiguity has been noted as a concern in food labels as well (e.g., Voordouw et al. 2009). Even though respondents with a need for allergen avoidance are motivated to find the right information, it is striking that also in food labelling respondents with food allergies showed a strong preference for labels including a symbol (DunnGalvin et al., 2019; Marra et al., 2017). Thus, even though not in line with the current regulation, depicting the allergen information with a visual cue as done in the Alternative 2 label is much more preferred by the target group than the Current label.

No general conclusion can be made whether prominent pictograms were more effective with consumers than text. The 'Children' icon,

which appeared much more prominently on the Alternative 2 label than on the other two tested options, without accompanying text, led to a twice as frequent recall of the message 'Keep away from children. Moreover, for the 'Eyes' hazard warning (which was conveyed using an equally sized icon on the Alternative 2 label, without text), the opposite effect was seen, with a two times higher recall for the Current label.

Some differences in the stickiness of specific types of hazards were seen between the three labels and the overall health hazard perception was somewhat stronger for the label containing the highest number of pictograms. Nevertheless, overall these differences were limited and cannot be considered as highly meaningful.

When respondents were asked to imagine an accident where someone in their household had splashed the product into the eyes, in every label condition, four in five respondents mentioned they would rinse the eyes with water. This is, indeed, the recommended action. However, nearly none of the respondents said they would consult the label in case of an accident. As such, these findings suggest that participants rely on intuition, prior knowledge and experience to determine how to safely use a product, rather than on the safety label. This is in line with the findings reported by ECHA (2012) that safety behaviours are influenced by an experience-related rather than an information-based hazard perception. It is also in line with prior studies that found label effectiveness to be lower when consumers are familiar with the product (Rogers et al., 2000).

In contrast to prior research on the effectiveness of pictorial information on safety labels (e.g., Andrews et al., 2014, Davis & Burton, 2016, Laughery & Wogalter, 2014; Niederdeppe et al., 2019; Sathar et al., 2016, Vigoroso et al., 2020; Wogalter et al., 2006), in a context of more complex, real-life labels tested in a group of target users, we could not find direct evidence of enhanced cognitive and behavioural effects of the more prominent use of pictograms on safety labels (except for the clarity of the allergen information). As such, previously reported results do not seem to generalize to a more complex real-life-like situation and thus we cannot provide a positive answer to RQ1 and RQ3. However, we could also not find any indication that the substantial simplification of the label's content had led to a poorer comprehension (as did Bagagiolo et al., 2019), to undesirable behavioural effects, or to a decreased overall hazard perception. Thus, in the context of a realistic label, the effectiveness of pictograms to convey hazard and safety information to the consumer seems to be at least equivalent to text.

Furthermore, there was evidence that the simplified labels were better liked. So, we did find a positive affective effect (RQ2). This is in line with Kalsher et al. (1996) who also found that labels with pictorials were preferred over and above labels without pictorials. Four out of five respondents preferred one of the simplified Alternative labels (with the strongest preference for Alternative 1), while only 17% preferred the most complex Current label. Thus, also for real-life labels, consumers preferred simpler graphical information rather than large amounts of text. As such, for safe use communication, the 'less is more' principle adhered to in marketing communications also seems applicable and prior results on the affective front seem to generalise to our more complex real-life-like situation.

Overall, this study highlights that the back labels of daily used consumer products, like detergents, are poorly effective at conveying safety messages. Irrespective of the label execution (current, more graphical, or highly simplified), none of the information elements were particularly well recalled and no really meaningful differences were detected. Nevertheless, the alternative executions were preferred by the respondents. Compared to the current text-rich label, these alternative labels did not bring about any decrease in the noticing, understanding and acting upon the safety instructions. Because of the higher preference, it is recommended that opportunities to increase label effectiveness be sought in the area of improved simplified graphical labels rather than in fine-tuning the current highly text-based approach. This research has implications for policy makers, marketers and consumer interest groups alike. Policy makers and marketers should prioritise consumer focused safety labels as a means to support effective hazard communication. That is, policy makers should seek to ensure that the regulatory framework enables marketers to provide labels that score high on attention-getting capacity, recall, liking and behavioural compliance. This should be tested in target groups that are most vulnerable, for example, people in need of allergen avoidance. This study showed that a simple revision such as adding the visual cue 'ALLERG' and highlighting only the allergens (rather than the full ingredient list) is already a step in the right direction. Furthermore, Alborzi et al. (2017) showed that lower educated groups do not pay much attention to detergent labels and are less likely (than higher educated groups) to take, for example, dosage instructions into account. Also, Miilunpalo and Räisänen (2019) find that only one in two consumers' course of action to determine detergent dosage is to consult the dosage instructions on the package and that many consumers administer excessive amounts of detergent. Improving the effectiveness of on pack labelling may in turn promote both safety and environmental benefits across the wider community and not just in higher educated groups. Consumer interest groups, in turn, could bring the complex labels to the attention of regulators (especially the difficult to comprehend allergen information) and request clear, less complex and easy to understand labels.

4.2 | Limitations and further research

A limitation of the current research is that the affective aspect—that is, label preference—was tested within subjects. This means that respondents were aware of the dimensions on which the labels differed (i.e., information presented in the form of text or visualized via pictograms). As such, the compromise effect may have occurred. That is, respondents may have voted for the Alternative 1 label as the 'middle' option because they did not want to lose on any of the two dimensions. Future research could test both Alternative labelling approaches one to one, to check whether the same preference of Alternative 1 over Alternative 2 holds. A related study limitation is that the labels were assessed for preference relative to the others, rather than being rated individually. While the development of an individual label

preference score was not within the scope of this work, it should be considered for future research.

As a compromise to not extend the length of the study beyond a reasonable attention span (which, based on experience, was assumed to be around 20 min), the hazard perception questions were not alternated with non-relevant fillers. This may potentially have biased the responses towards a higher hazard perception. However, as in practice, the responses were largely neutral, this is unlikely.

No actual behaviour was tested, although we believe that the 'safety intention' response gives a good picture of what would happen in real life in case of an accident. Nevertheless, future research should try to capture actual behaviour as well. Relatedly, we did not use a field study, but respondents participated in an online experimental study. As such, they were forced to look at the labels which likely induced non-natural behaviour. Our conclusions should, therefore, be interpreted with care. However, if any, we expect that consumers' attention in real life is even worse than what we observed here.

Finally, because the focus of the research was on the holistic effect of the labels rather than on the effect of individual label elements, the manipulations were not single-variable—but instead, entire labels were evaluated as integral concepts. Although our approach has the advantage of testing very realistic labels, taking into account concerns of regulation, marketers and consumers, the drawback is that we can only formulate conclusions on the overall design. It is recommended that future research explore the individual impact of the amount of information, amount of pictograms, pictogram size and the Allerg visual cue within a single alternative label concept. This may allow further improvement of the simplified graphical label executions, aiming to effectively increase the noticing, understanding and acting upon safety guidance compared to the current approach. Related to this, it would be interesting to use designs that use 2 × 2 designs in which the extent of textual information is manipulated together with the extent or size of pictorials. This would allow to investigate how text and visuals interact in their effects on cognitive, affective and behavioural measures. Moreover, in order to go beyond practical implications, future research could focus on investigating the impact of important psychological constructs that have been shown to have an impact on safety behaviour of consumers, such as perceived risk and self-efficacy. Relying on the Risk Perception Attitude framework (Rimal & Real, 2003), Lee and You (2020) classified chemical household product users into four groups according to the dimensions of perceived risk and self-efficacy and observed differential safety behaviour between the groups. Self-efficacy appeared to be important to lead consumers who perceived risk to engage in safe behaviour. It would, therefore, be interesting to investigate to what extent different labels affect perceived self-efficacy. Does dense information lead to confusion and overload of information (cf. Bialkova et al., 2013; SynapsesQuali, 2016) thereby reducing consumers' perceived self-efficacy? Do pictorials lead to a processing fluency experience (cf. Unkelbach, 2007; Winkielman et al., 2003) and hence enhance consumers' perceived self-efficacy? And when and for whom does self-efficacy then translate into compliant behaviour with the safety guidelines?

4.3 | Conclusions

To conclude, this study compared two simplified alternatives with a representative example of the currently used label of household detergent products, a regularly used low hazard product. Two main differences with prior research are that (a) the test and baseline labels in this research are more complex, as they did not only contain a safety warning but the safety warning was integrated in a label containing several other information elements and (b) the test population of this research consisted of target users who are familiar with the product type. Both factors could render the use of pictorials less effective than previously observed because both information clutter and product familiarity decrease consumers' attention. This is exactly what our results show. In our more complex, real-life-like case, the simplified labels containing less information and more pictorials (vs. the current more information-dense label), did not show any advantage on cognitive and behavioural measures, but not a disadvantage either. Importantly, they did have an advantage in terms of likeability. So, the previously reported positive results of pictorials seem to only partly generalise to more complex, real-life-like situations. Nevertheless, although mapping theoretical recommendations on a more complex real-life label does not immediately and straightforwardly result in a perfect label, moving further in the direction of simplified graphical labels—as advocated by theory—does seem promising. More research into how product labels can more effectively induce safe use behaviour is called for. Future research should examine for whom and under which circumstances graphics and text-based information are most effective at being noticed, understood, and acted upon in a complex product label. Ideally, the properties of both graphics (e.g., type, size, colour, etc.) and text (e.g., readability, size, font, etc.), and how they can be optimised should be considered in future research.

CONFLICTS OF INTEREST

The authors have declared no conflicts of interest for this article.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available via the A.I.S.E. website at [https://www.aise.eu/our-activities/regulatory-context/classification-labelling/better-regulation-safe-use.aspx].

ORCID

Maggie Geuens https://orcid.org/0000-0001-6535-7260
Dominic Byrne https://orcid.org/0000-0001-5729-6349
Bert Vandecasteele https://orcid.org/0000-0001-5145-7754

REFERENCES

Adams, A., Boersema, T., & Mijksenaar, M. (2010). Warning symbology: Difficult concepts may be successfully depicted with two-part signs. *Information Design Journal*, *18*(2), 94–106. https://doi.org/10.1075/idj.18.2.01ada

Alborzi, F., Schmitz, A., & Stamminger, R. (2017). Effects of socio-demographic factors on laundry behaviours in Europe and their implications on sustainability. *International Journal of Consumer Studies*, *41*(6), 671–684. https://doi.org/10.1111/ijcs.12380.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*(4), 569–576. https://doi.org/10.1037/0096-3445.136.4.569

Ancker, J. S., Senathirajah, Y., Kukafka, R., & Starren, J. B. (2006). Design features of graphs in health risk communication: A systematic review. *Journal of the American Medical Informatics Association*, *13*(6), 608–618. https://doi.org/10.1197/jamia.M2115

Andrews, J. C., Netemeyer, R. G., Kees, J., & Burton, S. (2014). How graphic visual health warnings affect young smokers' thoughts of quitting. *Journal of Marketing Research*, *51*(2), 165–183. https://doi.org/10.1509/jmr.13.0092.

Bagagiolo, G., Vigoroso, L., Caffaro, F., Micheletti Cremasco, M., & Cavallo, E. (2019). Conveying safety messages on agricultural machinery: The comprehension of safety pictorials in a group of migrant farmworkers in Italy. *International Journal of Environmental Research and Public Health*, *16*(21), 4180. https://doi.org/10.3390/ijerph16214180.

Bartels, M., Tillack, K., & Jordan Lin, C. T. (2018). Communicating nutrition information at the point of purchase: An eye-tracking study of shoppers at two grocery stores in the United States. *International Journal of Consumer Studies*, *42*(5), 557–565. https://doi.org/10.1111/ijcs.12474.

Bialkova, S., Grunert, K. G., & van Trijp, H. (2013). Standing out in the crowd: The effect of information clutter on consumer attention for front-of-pack nutrition labels. *Food Policy*, *41*, 65–74. https://doi.org/10.1016/j.foodpol.2013.04.010.

Boelhouwer, E., Davis, J., Franco-Watkins, A., Dorris, N., & Lungu, C. (2013). Comprehension of hazard communication: Effects of pictograms on safety data sheets and labels. *Journal of Safety Research*, *46*, 145–155. https://doi.org/10.1016/j.jsr.2013.06.001

Caffaro, F., Schmidt, S., Murphy, D. J., & Cavallo, E. (2018). Comprehension rates of safety pictorials affixed to agricultural machinery among Pennsylvania rural population. *Safety Science*, *103*, 162–171. https://doi.org/10.1016/j.ssci.2017.11.021

Chan, A. H. S., & Ng, A. W. Y. (2010). Investigation of guessability of industrial safety signs: Effects of prospective-user factors and cognitive sign features. *International Journal of Industrial Ergonomics*, *40*(6), 689–697. https://doi.org/10.1016/j.ergon.2010.05.002

Dalvie, M. A., Rother, H. A., & London, L. (2014). Chemical hazard communication comprehensibility in South Africa: Safety implications for the adoption of the globally harmonised system of classification and labeling of chemicals. *Safety Science*, *61*, 51–58. https://doi.org/10.1016/j.ssci.2013.07.013

Davis, C., & Burton, S. (2016). Understanding graphic pictorial warnings in advertising: A replication and extension. *Journal of Advertising*, *45*(1), 33–42. https://doi.org/10.1080/00913367.2015.1089426

De Pelsmacker, P., Geuens, M., & Van den Bergh, J. (2021). *Marketing communications* (7th ed.). Pearson Education (Financial Times/Prentice Hall).

DeJoy, D. M. (1989). Consumer product warnings: review and analysis of effectiveness research. *Proceedings of the Human Factors Society Annual Meeting*, *33*(15), 936–940. https://doi.org/10.1177/154193128903301508

Dörnyei, K. R., & Gyulavári, T. (2016). Why do not you read the label?–an integrated framework of consumer label information search. *International Journal of Consumer Studies*, *40*(1), 92–100. https://doi.org/10.1111/ijcs.12218.

Dowse, R., & Ehlers, M. S. (2005). Medicine labels incorporating pictograms: Do they influence understanding and adherence? *Patient Education and Counseling*, *58*(1), 63–70. https://doi.org/10.1016/j.pec.2004.06.012

DunnGalvin, A., Roberts, G., Regent, L., Austin, M., Kenna, F., Schnadt, S., Sanchez-Sanz, A., Hernandez, P., Hjorth, B., Fernandez-Rivas, M., Taylor, S., Baumert, J., Sheikh, A., Astley, S., Crevel, R., & Mills, C. (2019). Understanding how consumers with food allergies make decisions based on precautionary labelling. *Clinical & Experimental Allergy*, *49*(11), 1446–1454. https://doi.org/10.1111/cea.13479.

ECHA. (2012). *Communication on the safe use of chemicals: study on the communication of information to the general public (ECHA-12-A-01-EN)*. Author. https://echa.europa.eu/documents/10162/13559/clp_study_en.pdf

European Commission. (2011, May). *Special Eurobarometer 360. Consumer understanding of labels and the safe use of chemicals*. https://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_360_en.pdf

Fang, X., Singh, S., & Ahluwalia, R. (2007). An examination of different explanations for the mere exposure effect. *Journal of Consumer Research*, *34*(1), 97–103. https://doi.org/10.1086/513050

Friedmann, K. (1988). The effect of adding symbols to written warning labels on user behavior and recall. *Human Factors*, *30*(4), 507–515.

Handcock, H. E., Rogers, W. A., Schroeder, D., & Fisk, A. D. (2004). Safety symbol comprehension: Effects of symbol type, familiarity, and age. *Human Factors*, *46*(2), 183–195. https://doi.org/10.1518/hfes.46.2.183.37344.

Hawkins, S. A., & Hoch, S. J. (1992). Low-involvement learning: Memory without evaluation. *Journal of Consumer Research*, *19*, 212–225. https://doi.org/10.1086/209297

Hess, R., Visschers, V. H., Siegrist, M., & Keller, C. (2011). How do people perceive graphical risk communication? The role of subjective numeracy. *Journal of Risk Research*, *14*(1), 47–61. https://doi.org/10.1080/13669877.2010.488745.

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association*, *81*(396), 991–999. https://doi.org/10.1080/13669877.2010.488745

Houts, P. S., Doak, C. C., Doak, L. G., & Loscalzo, M. J. (2006). The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Education and Counseling*, *61*(2), 173–190. https://doi.org/10.1016/j.pec.2005.05.004

Kalsher, M. J., Wogalter, M. S., & Racicot, B. M. (1996). Pharmaceutical container labels: Enhancing preference perceptions with alternative designs and pictorials. *International Journal of Industrial Ergonomics*, *18*(1), 83–90. https://doi.org/10.1016/0169-8141(95)00033-X

Kaufmann, D., & Ramirez-Andreotta, M. D. (2019). Communicating the environmental health risk assessment process: Formative evaluation and increasing comprehension through visual design. *Journal of Risk Research*, *23*(9), 1177–1194. https://doi.org/10.1080/13669877.2019.1628098.

Kees, J., Burton, S., Andrews, J. C., & Kozup, J. (2010). Understanding how graphic pictorial warnings work on cigarette packaging. *Journal of Public Policy & Marketing*, *29*(2), 265–276. https://doi.org/10.1509/jppm.29.2.265.

Kelley, C. A., Gaidis, W. C., & Reingen, P. H. (1989). The use of vivid stimuli to enhance comprehension of the content of product warning messages. *Journal of Consumer Affairs*, *23*(2), 243–266. https://doi.org/10.1111/j.1745-6606.1989.tb00247.x.

Klaschka, U., & Rother, H. A. (2013). 'Read this and be safe!' Comparison of regulatory processes for communicating risks of personal care products to European and South African consumers. *Environmental Sciences Europe*, *2013*(25), 30. https://doi.org/10.1186/2190-4715-25-30

Laughery, K. R. (2006). Safety communications: Warnings. *Applied Ergonomics*, *37*(4), 467–478. https://doi.org/10.1016/j.apergo.2006.04.020

Laughery, K. R., & Wogalter, M. S. (2014). A three-stage model summarizes product warning and environmental sign research. *Safety Science*, *61*, 3–10. https://doi.org/10.1016/j.ssci.2011.02.012

Lee, M., & You, M. (2020). Safety Behaviors to reduce risk of using chemical household products: An application of the risk perception

attitude framework. *International Journal of Environmental Research and Public Health*, *17*(5), 1528.–https://doi.org/10.3390/ijerph17051528.

Lesch, M. F., Horrey, W. J., Wogalter, M. S., & Powell, W. R. (2011). Age-related differences in warning symbol comprehension and training effectiveness: Effects of familiarity, complexity, and comprehensibility. *Ergonomics*, *54*(10), 879–890. https://doi.org/10.1080/00140139.2011.606924

Marra, C. A., Harvard, S., Grubisic, M., Galo, J., Clarke, A., Elliott, S., & Lynd, L. D. (2017). Consumer preferences for food allergen labeling. *Allergy, Asthma & Clinical Immunology*, *13*(1), 19. https://doi.org/10.1186/s13223-017-0189-6.

Miilunpalo, S. M., & Räisänen, R. (2019). Clean laundry with pure conscience—A study on laundry practices among Finnish consumers. *International Journal of Consumer Studies*, *43*(2), 153–165. https://doi.org/10.1111/ijcs.12493.

Moreira, M. J., García-Díez, J., de Almeida, J. M. M. M., & Saraiva, C. (2019). Evaluation of food labelling usefulness for consumers. *International Journal of Consumer Studies*, *43*(4), 327–334. https://doi.org/10.1111/ijcs.12511.

Niederdeppe, J., Kemp, D., Jesch, E., Scolere, L., Greiner Safi, A., Porticella, N., Avery, R. J., Dorf, M. C., Mathios, A. D., & Byrne, S. (2019). Using graphic warning labels to counter effects of social cues and brand imagery in cigarette advertising. *Health Education Research*, *34*(1), 38–49. https://doi.org/10.1093/her/cyy039.

Noar, S. M., Hall, M. G., Francis, D. B., Ribisl, K. M., Pepper, J. K., & Brewer, N. T. (2016). Pictorial cigarette pack warnings: A meta-analysis of experimental studies. *Tobacco Control*, *25*(3), 341–354. https://doi.org/10.1136/tobaccocontrol-2014-051978.

Okan, Y., Stone, E. R., & Bruine de Bruin, W. (2018). Designing graphs that promote both risk understanding and behavior change. *Risk Analysis*, *38*(5), 929–946. https://doi.org/10.1111/risa.12895

Okan, Y., Stone, E. R., Parillo, J., Bruine de Bruin, W., & Parker, A. M. (2020). Probability size matters: The effect of foreground-only versus foreground+ background graphs on risk aversion diminishes with larger probabilities. *Risk Analysis*, *40*(4), 771–788. https://doi.org/10.1111/risa.13431.

European Commission. (2012). *Report from the Commission to the European Parliament and the Council on communication on the safe use of chemicals* (Ref. COM(2012) 630 final). https://ec.europa.eu/transparency/regdoc/rep/1/2012/EN/1-2012-630-EN-F1-1.pdf

Rhoades, T. P., Frantz, J. P., & Miller, J. M. (1990). Emerging methodologies for the assessment of safety related product communications. *Proceedings of the Human Factors Society Annual Meeting*, *34*(14), 998–1002. https://doi.org/10.1177/154193129003401408

Rimal, R. N., & Real, K. (2003). Perceived risk and efficacy beliefs as motivators of change: Use of the risk perception attitude (RPA) framework to understand health behaviors. *Human Communication Research*, *29*(3), 370–399. https://doi.org/10.1111/j.1468-2958.2003.tb00844.x.

Rogers, W. A., Lamson, N., & Rousseau, G. K. (2000). Warning research: An integrative perspective. *Human Factors*, *42*(1), 102–139. https://doi.org/10.1518/001872000779656624.

Roose, G., Vermeir, I., Geuens, M., & Van Kerckhove, A. (2019). A match made in heaven or down under? The effectiveness of matching visual and verbal horizons in advertising. *Journal of Consumer Psychology*, *29*(3), 411–427. https://doi.org/10.1002/jcpy.1088.

Sathar, F., Dalvie, M. A., & Rother, H. A. (2016). Review of the literature on determinants of chemical hazard information recall among workers and consumers. *International Journal of Environmental Research and Public Health*, *13*(6), 546.–https://doi.org/10.3390/ijerph13060546.

Scazzola, R., Boeije, G., Bøtker Pedersen, E., Brenneman, W., Cagáňová, B., Celentano, A., De Coninck, E., Desel, H., Ebbehøj, N., Färber, E., Sesana, F., & Zacharov, S. (2019). Eye hazard classification according to UN GHS/EU CLP and the severity of eye symptoms caused by accidental exposures to detergents and cleaning products. *Regulatory Toxicology and Pharmacology*, *105*(2019), 69–76. https://doi.org/10.1016/j.yrtph.2019.04.004

Siegrist, M., Cousin, M. E., & Keller, C. (2008). Risk communication, prenatal screening, and prenatal diagnosis: The illusion of informed decision-making. *Journal of Risk Research*, *11*(1–2), 87–97. https://doi.org/10.1080/13669870701574015

Stone, E. R., Bruine de Bruin, W., Wilkins, A. M., Boker, E. M., & MacDonald Gibson, J. (2017). Designing graphs to communicate risks: Understanding how the choice of graphical format influences decision making. *Risk Analysis*, *37*(4), 612–628. https://doi.org/10.1111/risa.12660

Stone, E. R., Gabard, A. R., Groves, A. E., & Lipkus, I. M. (2015). Effects of numerical versus foreground-only icon displays on understanding of risk magnitudes. *Journal of Health Communication*, *20*(10), 1230–1241. https://doi.org/10.1080/10810730.2015.1018594

SynapsesQuali. (2016, July). *Safety information on household products. Better Regulation & Safe Use Project (A.I.S.E.) Research conducted in Belgium, Poland & Spain*. www.aise.eu

Ta, G. C., Mokhtar, M. B., Mokhtar, H. A. B. M., Ismail, A. B., & Yazid, M. F. B. H. A. (2010). Analysis of the comprehensibility of chemical hazard communication tools at the industrial workplace. *Industrial Health*, *48*(6), 835–844. https://doi.org/10.2486/indhealth.MS1153

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33*(1), 219–230. https://doi.org/10.1037/0278-7393.33.1.219.

Vigoroso, L., Caffaro, F., & Cavallo, E. (2020). Occupational safety and visual communication: User-centred design of safety training material for migrant farmworkers in Italy. *Safety Science*, *121*, 562–572. https://doi.org/10.1016/j.ssci.2018.10.029.

Voordouw, J., Cornelisse-Vermaat, J. R., Yiakoumaki, V., Theodoridis, G., Chryssochoidis, G., & Frewer, L. J. (2009). Food allergic consumers' preferences for labelling practices: A qualitative study in a real shopping environment. *International Journal of Consumer Studies*, *33*(1), 94–102. https://doi.org/10.1111/j.1470-6431.2008.00735.x.

Wilkinson, R. L., Cary, J. W., Barr, N. F., & Reynolds, J. (1997). Comprehension of pesticide safety information: Effects of pictorial and textual warnings. *International Journal of Pest Management*, *43*(3), 239–245. https://doi.org/10.1080/096708797228744

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: implications for evaluative judgment. In: J. Musch, & K. C. Klauer (Eds.), *The psychology of evaluation: affective processes in cognition and emotion* (pp. 189–217). : Erlbaum. https://doi.org/10.4324/9781410606853.

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health campaigns. *Health Education & Behavior*, *27*(5), 591–615. https://doi.org/10.1177/109019810002700506.

Wogalter, M. S., Silver, N. C., Leonard, S. D., & Zaikina, H. (2006). Warning symbols. In M. S. Wogalter (Ed.). *Handbook of warnings* (pp. 159–176). : Erlbaum. https://doi.org/10.1201/9781482289688.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Geuens M, Byrne D, Boeije G, Peeters V, Vandecasteele B. Investigating the effectiveness of simplified labels for safe use communication: The case of household detergents. *Int J Consum Stud*. 2021;45:1410–1424. https://doi.org/10.1111/ijcs.12662