Memory of Medical Scenarios for End-of-Life Support Preferences

Elvira García-Bajos and Malen Migueles

Universidad del País Vasco (Spain)

Abstract. We studied memory of health scenarios for end-of-life support decisions and stability of life support preferences. Psychology students (n = 36, age M = 27.25, SD = 6.21, 64% females) were administered the Life Support Preferences Questionnaire to assess their memory of six medical scenarios with different prognoses, care treatment, and end-of-life support choices. Recall, recognition, confidence and decision stability were assessed immediately and one month later. Correct recall decreased and incorrect recall increased from immediate to long-term recall, F(2, 68) = 74.38, p < .001, η²p = .69. In recall, participants spontaneously gave false information consistent with prior knowledge of illnesses and medical scenarios. Participants who had suffered a disease or serious accident did worse on correct recall, F(1, 34) = 6.59, p = .015, η²p = .16, and had more errors, F(1, 34) = 4.68, p = .038, η²p = .12, than participants who had not. In the recognition test there were no differences between hits and false alarms, showing the difficulty in discriminating between true and false contents. Confidence was greater for hits than for false alarms, F(1, 34) = 10.86, p = .002, η²p = .24, but this subjective measure did not seem to be a good predictor of accuracy because confidence was greater than the mean value for hits (p = .001, d = 1.74) and for false alarms (p = .001, d = 0.96). Long-term memory was quite poor and biased, but life support preferences did not change much.

Keywords: end-of-life, medical scenarios, memory, negative events, support preferences.

Our life experience is made up of positive and negative autobiographical events (Berntsen, 2002; García-Bajos & Migueles, 2013). Accidents and illnesses, either our own or those of loved ones, mark our worries and fears in the different stages of our lives. Negative autobiographical thoughts about the future include the death of people close to us and our own death, especially the circumstances in which we die (Allmark, 2002). People are aware of their fear of suffering, prolonged pain, cognitive impairment to make decisions or a loss of consciousness (Conway, Singer, & Tagini, 2004). For these reasons, many people choose to execute Advance Directives (ADs) on end-of-life care and thus prevent painful circumstances for oneself or one’s loved ones (van Wijmen, Pasman, Widdershoven, & Onwuteaka-Philipsen, 2014).

Stability and false memories for end-of-life decisions
People use ADs to provide a legal and medical framework for the life-sustaining treatments they wish to receive at the end of life. We wanted to see whether medical decisions made when people are younger or long before the onset of a serious illness reflect current decisions. Some studies show changes in life-sustaining treatment preferences (Ditto et al., 2003) while in others up to 86% of decisions remain stable 18 months later (Barrio-Cantalejo et al., 2013). Preferences over time and changes in health status show that preference stability is generally greater among inpatients and seriously ill outpatients than among older adults without serious illnesses (Auriemma et al., 2014). The studies show that people have little awareness of changes in life support preferences (Gready et al., 2000; Sharman, Garry, Jacobson, Loftus, & Ditto, 2008), either because they do not remember which support preferences they chose at the time or because they changed them to meet their current needs. Sharman et al. (2008) used the Life Support Preferences Questionnaire by Ditto et al. (2001), which includes nine common end-of-life scenarios, in two sessions 12 months apart for older adults, and four months apart for younger adults. They found changes in 23% of the decisions both in the young and in the older adults. Changes were more likely to be from wanting treatment to not wanting treatment than the reverse in the older adults, while...
younger participants were no more likely to change from wanting treatment to not wanting treatment than the reverse. More interestingly, 75% of the participants (69% of the young adults) who changed their decisions thought that they had not. Gready et al. (2000) found similar results, with 80% of the changes undetected. We should point out that when the participants were asked to provide subjective measures of metamemory, such as response confidence, they rated their correct and incorrect memories with the same confidence. Thus, in addition to producing false memories for life-sustaining treatment decisions, participants perceive false memories as if they were real. An interesting aspect is that mood can affect the decisions. Participants in negative moods more often change their current decisions and falsely remember their previous decisions than participants in positive moods (Sharman, 2011).

**Memory of medical scenarios in the Life Support Preferences Questionnaire**

Studies have been conducted on the memory of treatment preferences in situations involving end-of-life decisions (Sharman, 2011; Sharman et al., 2008), but data is lacking on how medical scenarios and treatment alternatives in the Life Support Preferences Questionnaire (LSPQ) are remembered in the long term (Barrio-Cantalejo et al., 2008; Beland & Froman, 1995). We questioned whether support preferences results are consistent with the memory of medical scenarios and treatments. While the importance of mood has been pointed out in the memory for end-of-life treatment decisions (Sharman, 2011), we do not know whether this type of emotionally charged medical information is correctly or incorrectly remembered over time. Understanding the memory of medical scenarios and treatments can help explain the cases in which decisions are relatively stable and the characteristics of the clinical situations more inclined to changes and distortions of memory. We were also interested in knowing whether variables such as gender, depression or fear of death affect performance. Depression is often associated with physical illness (Ruttley & Reid, 2006) and with fear of death (Stenzel, Vaske, Kühl, Kenn, & Rief, 2015), and gender differences in diseases and health care (e.g., Denton, Prus, & Walters, 2004) can modulate memory and life-support preferences.

We do not have prior data on the memory for scenarios and treatments for end-of-life preferences, but we would expect memory for this type of information to be much like memory for medical information. When relevant decisions must be made, patients receive information on their illness, treatment options and consequences. But how much of this information do they remember? Between 40–80% of medical information is forgotten immediately; one month later recall drops to 12.8%, and the more information a doctor provides the worse the patients’ memory (McGuire, 1996). For example, Brown and Park (2002) observed that for familiar illnesses (e.g., breast cancer) participants were only able to remember between 6% and 13% of new information provided to them about the disease. In general, we have better memory for information consistent with our knowledge, but false or inaccurate information is also incorporated into our memory (Kessels, 2003). For example, Godwin (2000) examined female patients’ recall of information related to reduction mammoplasty. On as many as three occasions the patients were explained the characteristics, complications and consequences of the procedure. The women were very satisfied with the information process, which included drawings, diagrams and photographs, but at six days they only retained an average of three out of 12 possible postoperative complications. Interestingly, performance was even worse when there were postsurgery complications.

Aspects such as complicated medical terms or patient expectations also contribute to poor memory for medical information (Ley, 1979). Moreover, stress from negative emotional contents and news that threatens our lives or personal safety can shift our attention to more central and relevant information (e.g., the diagnosis), leaving few mental resources available to attend to other specific content and information (Kessels, 2003). In any event, medical content generally involves disturbing or negative information that we do not want to remember. Recent research shows that we tend to block or forget negative events (García-Bajos & Migueles, 2017; García-Bajos, Migueles, & Aizpurua, 2017; Szpunar, Addis, & Schacter, 2012).

**Objectives and hypotheses**

Unlike previous studies focusing on end-of-life preference decisions, our aim was to examine in a single study both the memory for medical scenarios in the LSPQ and the stability of treatment decisions and preferences. To evaluate memory immediately and at one month we used a cued-recall test and a true/false recognition test. The cued-recall test allowed us to steer recall toward specific clinical situations to examine memory accuracy and the nature of errors, while the recognition test was used to assess whether participants were able to differentiate between correct and erroneous information. We expected poorer performance in recall than in recognition. Recall requires the retrieval of information specific to each medical situation and treatment, and it is known that emotional content, especially when negative, is remembered in a more general, less detailed fashion than positive content.
(Szpunar et al., 2012). Recognition, on the other hand, only requires participants to verify whether the information in the recognition test corresponds to what they remember from the LSPQ. A second objective was to further understand the stability of life-support preferences. We specifically sought to examine whether there are differences between medical cases based on life situation, treatments and consequences, and if there are differences based on whether a person has experienced a serious accident or illness. An additional objective was to determine the effects of gender differences on memory for negative information.

In sum, the theoretical interest of this study was to investigate cognitive factors and medical background affecting memory and metamemory of scenarios for end-of-life support preferences. From an applied perspective, another important issue concerns the effects of psychological factors such as death anxiety or depression in advance directive decisions.

Method

Participants

Thirty-six psychology students aged 22 to 40 ($M = 27.25$ years; $SD = 6.21$), 23 women and 13 men, volunteered to participate in this experiment. All interactions with participants followed the ethical principles and code of conduct of psychologists. The study was carried out in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the University of the Basque Country. Before starting the experiment, written informed consent was obtained from all the participants. Owing to the personal and emotional dimension of the study, the participants were explicitly told that they could drop out of the experiment at any time.

To confirm that the study had sufficient statistical power, an analysis using the G*POWER software (Faul, Erdfelder, Lang, & Buchner, 2007) was applied to determine how many participants would be required to achieve a power of 0.95 with an alpha of .05 to detect a medium effect size of $f = 0.25$. A one-group design, with two memory tasks (recall, recognition) and two evaluation intervals (immediate, 1-month) as within-participants measures would require 36 participants, that is, the sample size tested in the present study.

Materials

To study end-of-life support preferences and memory of ADs, we used the LSPQ (Beland & Froman, 1995) validated for the Spanish population (Barrio-Cantalejo et al., 2008). It describes six different health states: (1) coma state, (2) vegetative state, (3) cancer, (4) Alzheimer’s disease, (5) constant pain, and (6) kidney failure. Each case includes a description of the clinical situation, the current medical scenario and two alternatives: Accept or refuse medical treatment, together with the consequences of each option, allowing participants to choose one of the two.

To examine memory of the LSPQ, two versions were designed, each containing three medical cases, to counterbalance immediate and one-month recall, in order to control for memory carryover effects due to repeated evaluations of the same medical cases. Several aspects were taken into account when drafting the two versions. To control for primacy and recency effects, cases 1 and 6 (coma state, kidney failure) of the questionnaire were included in different subgroups. Cases 3 and 5 (cancer, constant pain), which involve situations that allow patients to make their own decisions, were assigned to different subgroups. Cases 2 and 4 (vegetative state, Alzheimer’s disease), which would require third-party treatment decisions, were also assigned to different subgroups. The severity of the medical situation and the number of units of information were also taken into account. Using these aspects, two versions of the questionnaire were developed, one with cases 1, 4, and 5 (coma state, Alzheimer’s disease and constant pain) and another with cases 2, 3 and 6 (vegetative state, cancer and kidney failure). To evaluate recall, in each of the two versions of the recall questionnaire a total of 20 units of information were identified, corresponding to the clinical situations and current medical scenario (e.g., Case 2, vegetative state; the patient has pneumonia), treatments and consequences (e.g., if not treated with antibiotics, the patient will die). Each clinical case was presented in a response notebook, followed by a blank space in which participants wrote down all the information they could remember about the clinical situation, treatments and consequences.

Two versions were also designed for the true/false recognition test, one with cases 1, 4, and 5 (coma state, Alzheimer’s disease and constant pain) and another with cases 2, 3 and 6 (vegetative state, cancer and kidney failure). Each recognition test contained 18 sentences, nine true sentences (e.g., for cancer, I choose to refuse treatment even if it shortens my life) and nine false sentences (e.g., for the same clinical case of cancer, I choose to accept radiation therapy to prolong my life even though it causes side effects). Both versions of the recognition test included clinical situations, treatments and consequences of the three cases.

Design

A within-participants experimental design was used. The participants responded to the LSPQ immediately and at one month. Recall and recognition of the contents in the LSPQ were also evaluated immediately and at one month.
Procedure

The experiment was conducted either individually or in small groups of no more than four. Participants were told that the experiment was part of a study on advance directives for end-of-life health care decisions. Participants completed the LSPQ at the beginning of the first session before the memory tests, and at the end of the second session after the memory tests.

At the initial session, the participants first completed the consent form. They then were asked to fill in a form with their personal details and answer questions about the ADs document: Whether they had ever heard of an AD, if they would consider signing an AD, or if they already have signed an AD document. They were also asked if someone close to them had died or suffered a serious illness or if they themselves had suffered a serious illness or accident. Lastly, they rated their current state of health on a scale of 1 (very poor) to 5 (very good). The participants then read and answered the LSPQ.

After a five-minute distractor task, they completed the first recall test, followed by a recognition test involving three clinical cases, treatments and consequences. For the immediate and long-term recall and recognition, two versions were counterbalanced, each one containing three of the six cases in the LSPQ. For each recall test the names of the three cases were written on separate sheets of paper. Participants were asked to write down everything they could remember about the three clinical situations, treatments and consequences. They were given enough time to remember each case. The same cases were examined in the recognition test. The participants had to respond to 18 phrases about the three clinical cases, say whether they were true or false, and rate their level of confidence in each answer they provided on a scale of 1 (not at all certain) to 7 (totally certain).

The second experimental session took place one month later. First, the participants completed the recall and recognition tests for the three clinical cases in the LSPQ that had not been examined at the first session. Then they responded again to the six medical cases in the LSPQ that had not been examined at the first session. Omissions in recall were also analyzed.

Recall of the medical scenarios in the LSPQ

Recall was corrected by assigning one point for every specific piece of reported information from the LSPQ. The 20 units of information from the three clinical cases in the immediate evaluation and the 20 from the three clinical cases in the one-month evaluation were used to assess recall. The two checklists of 20 units each created for Cases 1, 4, 5 (coma, Alzheimer’s disease and constant pain) and Cases 2, 3, 6 (vegetative state, cancer and kidney failure) also served to assess possible errors. We found two types of errors in recall: Gist errors in the form of remembering information consistent with the case but not present in the LSPQ, which they added on their own (e.g., indicate weight loss or lack of appetite as a consequence of chemotherapy); and source errors, when the participants remembered information from the document but mistakenly attributed it to the wrong clinical case or treatment (e.g., indicate pneumonia, which corresponds to the clinical case vegetative state, and incorrectly associate with Alzheimer’s disease). Omissions in recall were also analyzed.
Using the correct answers and errors, a 3 x 2 (recall: immediate, 1-month) x 2 (gender: women, men) ANOVA was conducted. Post hoc comparisons for the recall factor and significant interactions were performed using the Bonferroni adjustment. The factors recall, F(2, 68) = 49.07, p < .001, η²p = .59, and interval, F(1, 34) = 72.65, p < .001, η²p = .68, were significant, as was the recall x interval interaction, F(2, 68) = 74.38, p < .001, η²p = .69. The participants remembered more correct units (M = 5.89) than gist errors (M = 1.58) or source errors (M = 1.90), with no significant differences between gist and source errors. More units were remembered immediately (M = 13.09) than at one month (M = 5.89). Errors increased significantly from immediate to long-term recall (2.95–4.25 = +1.30), t(35) = –2.14, p < .04, d = 0.48, but a surprising amount of information was lost from the immediate to the one-month evaluation in terms of correct recall (10.14–4.25 = −5.89), t(35) = 12.56, p < .001, d = 2.61 (Figure 1). Thus, in the Bonferroni post-hoc comparisons for the interaction recall (correct, incorrect) x interval (immediate, 1-month) we found a higher average of correct than incorrect units (10.14 vs. 2.94, p < .001) immediately, and greater incorrect than correct recall (4.25 vs. 1.64, p < .001) at one month. The gender factor was not significant, nor were the interactions of gender with correct and incorrect recall or with interval.

To analyze omissions in recall, a 2 x 2 (interval: Immediate, 1-month) x 2 (gender: women, men) ANOVA was conducted. The factor interval, F(1, 34) = 72.65, p < .001, η²p = .68, was significant. There were more omissions at one month (M = 14.11) than immediately (M = 6.92).

Neither the factor gender nor the interaction interval x gender were significant.

Using the correct and incorrect responses, we calculated recall accuracy by dividing the correct units recalled by the sum of correct and incorrect units recalled. A 2 x 2 (interval: Immediate, 1-month) x 2 (gender: women, men) mixed ANOVA was used to analyze accuracy. Only the interval factor was significant, F(1, 34) = 82.02, p < .001, η²p = .72. Accuracy was greater immediately (M = .76, SD = .17) than at 1-month (M = .25, SD = .30). Recall accuracy was significantly higher than .5 (chance level) immediately, t(35) = 9.11, p < .001, d = 1.52, and below chance performance at 1-month, t(35) = –4.86, p < .001, d = .81.

Another consideration in recall was whether the participants had experienced the death or illness of someone close to them and whether they had personally suffered a serious illness or accident (Table 2). For each of these criteria we conducted two 2 x 2 (group: Experienced/suffered, not experienced/suffered) x 2 (recall interval: Immediate, 1-month) ANOVAs, one for correct recall and another for errors.

There were no differences in recall of the LSPQ in either correct responses or errors based on whether participants had experienced a serious illness or death of a person close to them. However, there were significant differences in recall between the participants who had personally suffered an important illness or serious accident and those who had not. In the correct recall the factors group, F(1, 34) = 6.59, p = .015, η²p = .16, and interval, F(1, 34) = 105.45, p < .001, η²p = .76, and the group x interval interaction, F(1, 34) = 16.15, p < .001, η²p = .32, were significant. Recall was greater for participants who had not suffered (M = 6.43) than for those who had suffered a serious illness or accident (M = 4.00), but the differences were significant only in immediate recall (11.29 vs. 6.13), t(34) = 3.70, p < .01, d = 1.48, not at one month (1.57, 1.88). In the errors, only the group factor, F(1, 34) = 4.68, p = .038, η²p = .12, was significant. The participants who had suffered serious situations had more errors (M = 4.88) than those who had not (M = 3.23).

**Recognition of the medical scenarios in the LSPQ**

For recognition of the LSPQ contents, we counted the hits (answer true to true sentences in the document) and the false alarms (answer true to false sentences) (Table 3). In the ANOVA 2 (answers: Hits, false alarms)
In summary, like the recall test, the analyses indicate that recognition performance was poor and that participants had a high rate of false alarms, demonstrating their difficulty in discriminating between true and false contents in the LSPQ.

In evaluating recognition consideration was given to whether participants had experienced a serious illness or death of someone close to them or if they themselves had suffered a serious illness or accident. For each of these criteria, experienced and personally suffered, we applied four ANOVAs 2 (group: Experienced/suffered, not experienced/suffered) x 2 (recall interval: Immediate, 1-month) to analyze their effects in hits, false alarms, $A'$ accuracy scores and $B^D$ scores for response bias in recognition. There were no significant differences between having experienced/not experienced or suffered/not suffered serious situations in any of the recognition measurements.

Confidence in recognition

The participants rated their confidence in recognition from 1 (no certainty) to 7 (total certainty). Average men) ANOVAs were performed. In accuracy, only the interval factor was significant, $F(1, 34) = 5.99, p = .02$, $\eta^2_p = .15$. Accuracy in recognition was better immediately than a month later (Table 3). When comparing the accuracy values immediately and at one month with .5 probability (chance performance), accuracy was greater than chance in the immediate evaluation, $t(35) = 2.44, p = .02, d = 0.41$, but not at one month, $t(35) = .06, p = .95$. However, there were no significant changes in response criteria from the immediate to the long-term evaluation. Participants adopted a liberal or risky response criterion, with a tendency to say yes/true in the immediate and one-month recognition tests. When comparing the scores with the value 0 (neutral response criterion), significant differences were observed both immediately, $t(35) = –2.44, p = .02, d = 0.41$, and at one month, $t(35) = –4.11, p < .001, d = 0.68$. In summary, like the recall test, the analyses indicate that recognition performance was poor and that participants had a high rate of false alarms, demonstrating their difficulty in discriminating between true and false contents in the LSPQ.
conﬁdence scores are shown in Table 3. To analyze conﬁdence, a 2 (responses: Hits, false alarms) x 2 (interval: Immediate, 1-month) x 2 (gender: Women, men) ANOVA was performed. The response, F(1, 34) = 10.86, p = .002, $\eta^2_p = .24$, and the interval, F(1, 34) = 37.21, p < .001, $\eta^2_p = .52$, factors were signiﬁcant. Conﬁdence was greater in hits (M = 5.28) than in false alarms (M = 4.86) and in the immediate evaluation (M = 5.66) than at one month (M = 4.48). Comparing these values with the mean score of 4, we found greater conﬁdence than 4 in hits, t(35) = 10.42, p = .001, d = 1.74, and false alarms, t(35) = –5.79, p = .001, d = 0.96 immediately, t(35) = 10.99, p = .001, d = 1.83, and at one month, t(35) = 3.16, p = .01, d = 0.53. Although the participants showed a certain amount of control over their responses, i.e. greater conﬁdence in hits than in false alarms, conﬁdence was not a good predictor of accuracy since they also rated their false alarms with high levels of conﬁdence.

Another factor taken into account in recognition was whether or not participants had experienced a serious illness or death of someone close to them or had personally suffered a serious illness or accident. For each of these criteria we applied two 2 (group: Experienced/suffered, not experienced/suffered) x 2 (recall interval: Immediate, 1-month) ANOVAs to analyze their effects on conﬁdence in hits and false alarms. There were no signiﬁcant differences in conﬁdence based on whether participants had experienced or personally suffered serious illnesses or accidents.

Stability of life support preferences

The participants in this study answered the LSPQ on two occasions. This made it possible to examine whether there were ﬂuctuations in their choices. That is, if there were changes in their life support preferences from the immediate evaluation to the evaluation conducted one month later. The results are presented in Table 4. There was an average of 82.33% stability of life support preferences, and immediate and one-month comparisons using the McNemar test showed no signiﬁcant changes from immediate to one-month evaluation for any of the clinical cases (p > .05 in all comparisons).

Regarding the participants’ decisions to accept or refuse treatment, they signiﬁcantly chose to refuse treatment in Case 1 (coma state) both immediately, $\chi^2(1) = 11.11$, p < .01, $w = 0.79$, and at one month, $\chi^2(1) = 7.11$, p < .01, $w = 0.63$, and in Case 2 (vegetative state) immediately, $\chi^2(1) = 16$, p < .001, $w = 0.94$, and at one month, $\chi^2(1) = 7.11$, p < .01, $w = 0.63$. However, they accepted the treatments in Case 3 (cancer) immediately, $\chi^2(1) = 13.44$, p < .001, $w = 0.86$, and at one month, $\chi^2(1) = 5.44$, p = .02, $w = 0.55$, Case 5 (constant pain) immediately, $\chi^2(1) = 11.11$, p < .01, $w = 0.79$, and at one month, $\chi^2(1) = 4$, p < .05, $w = 0.47$, and Case 6 (kidney failure) immediately, $\chi^2(1) = 13.44$, p < .001, $w = 0.86$, and at one month, $\chi^2(1) = 7.11$, p < .01, $w = 0.63$. There were no differences between acceptances and refusals of treatment in Case 4 (Alzheimer’s) either immediately or at one month.

Cases 1 and 2, in which most participants refused treatment, were analyzed separately from cases 3 and 6, in which the majority did accept treatment, and from cases 4 and 5, which showed fewer differences between acceptance and refusal of treatment. In Case 1, the participant is in a coma and on a breathing machine following a stroke; would they want cardiopulmonary resuscitation if their heart stopped beating? In Case 2, the participant is an elderly single person in a vegetative state without hope of recovery; If they had pneumonia would they want to be treated with antibiotics or be allowed to die? The treatments in cases 1 and 2 would do nothing to alter the coma or vegetative state, and the participants chose to be allowed to die. In cases 3 and 6, the treatments were chemotherapy for cancer and being connected to a dialysis machine for kidney failure. Here the participants preferentially chose treatment over being allowed to die. Lastly, there was less agreement in cases 4 and 5, in which participants had to choose whether they would have a leg amputated if they had Alzheimer’s disease and developed an infection, or if they would want to be provided nutrition with a feeding tube if they suffered from constant chronic pain. It appears that participants opted for treatment in cases where they there was a glimmer of hope for a cure or improvement, but when the prognosis was very negative with no chance of getting better, people tended to refuse treatment.

In terms of gender, the only difference was observed in Case 3 (cancer), in the one-month evaluation, $\chi^2(1) = 5.20$, p = .023, $w = 0.54$, where there was a higher percentage of women (53%) than men (17%) who accepted chemotherapy, allowing them to live longer, even though it would make them feel sick, and there were more men (19%) than women (11%) who refused treatment. In all the other cases, treatment acceptance and refusal was balanced between men and women.

We also analyzed any possible differences based on whether participants had experienced a serious illness or death of a person close to them or had themselves suffered a serious illness or accident in their decisions to accept or refuse treatment for cases 1 (coma), 2 (vegetative state), 3 (cancer), 4 (Alzheimer’s), 5 (constant pain) and 6 (kidney failure), in the immediate evaluation and at one month. We only found signiﬁcant differences between acceptance and refusal of treatments in the one-month evaluation and only for two cases. There were signiﬁcant differences in whether they had...
experienced a serious illness or death of someone close to them in Case 3, cancer, $\chi^2(1) = 6.84$, $p = .009$, $w = 0.62$, and in whether they had personally suffered a serious illness or accident in Case 5, constant pain, $\chi^2(1) = 8.04$, $p = .005$, $w = 0.67$.

In Case 3 (cancer), of the 69.44% of participants who chose chemotherapy, the percentage was higher for participants who had experienced a serious illness or death of someone close to them (63.88%) than those who had not (5.55%). But there were no significant differences in participants who refused chemotherapy (30.55%) between those who had (16.66%) or had not (13.88%) experienced a serious illness or death of someone close to them. The experience of successful cancer treatment (e.g., 80% survival rate for breast cancer; Chirlaque et al., 2010) and the personal experience of seeing friends and loved ones survive the disease may have furthered the tendency to accept chemotherapy treatment. In Case 5 (constant pain), of the 66.66% of participants who chose to accept artificial nutrition and hydration, the percentage was higher for those who had not personally suffered a serious illness or accident than for those who had (63.88%) vs. 5.55%), whereas among those who had refused treatment (33.33%) there were no significant differences among participants who had personally suffered a serious illness or accident (16.66%) and those who had not (16.66%). In the case of constant pain, the people who had not personally suffered a serious illness or accident may have decided to accept artificial nutrition and hydration owing to inexperience in this type of problem and a lack of knowledge about treatments and implications.

### Table 4. Stability of Life Support Preferences. Percentage of Participants who Accepted the Medical Treatment for the Six Cases of the Life Support Preferences Questionnaire (LSPQ)

<table>
<thead>
<tr>
<th>Clinical cases and treatments</th>
<th>First</th>
<th>1-Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1: Stroke resulting in coma, unable to breathe without a machine</td>
<td>22.2%</td>
<td>19.4%</td>
</tr>
<tr>
<td>Scenario: Heart stops beating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment: Cardiopulmonary resuscitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 2: Elderly single person in persistent vegetative state</td>
<td>16.7%</td>
<td>27.8%</td>
</tr>
<tr>
<td>Scenario: Pneumonia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment: Antibiotics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 3: Diagnosed with cancer that probably cannot be cured</td>
<td>80.6%</td>
<td>69.4%</td>
</tr>
<tr>
<td>Scenario: Side effects of chemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment: Chemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 4: Advanced Alzheimer’s disease and diabetes</td>
<td>58.3%</td>
<td>63.9%</td>
</tr>
<tr>
<td>Scenario: Developed an infection in the leg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment: Amputation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 5: Constant untreated hip pain. Severely withdrawn</td>
<td>77.8%</td>
<td>66.7%</td>
</tr>
<tr>
<td>Scenario: Refuses to eat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment: Artificial nutrition and hydration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 6: Adolescent suffered trauma at birth has poor quality of life</td>
<td>80.6%</td>
<td>72.2%</td>
</tr>
<tr>
<td>Scenario: Kidney failure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment: Daily dialysis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Cases, scenarios and treatments based on the LSPQ (Barrio-Cantalejo et al., 2008; Beland & Froman, 1995).

Death anxiety, depression, and life support preferences

We examined whether factors such as death anxiety or level of depression affect life support preferences and memory for medical scenarios in the LSPQ.

To evaluate death anxiety, we used the 15-item Death Anxiety Scale (Templer, 1970). This test differentiates between no anxiety (0–4), mild anxiety (5–7), moderate anxiety (8–11) and severe anxiety (12–15). The mean score was 6.28 ($SD = 2.67$), assessed as mild death anxiety. When participants were divided into two groups, with death anxiety scores below or above 8 points, there were no significant differences among participants who had personally suffered a serious illness or death of someone close to them.

To evaluate the level of depression, we used the Beck Depression Inventory–II (Beck et al., 1996). The mean score was 7.11 (range 0–17; $SD = 4.19$) for the 21 test items, in the range of mild depression (mild 0–13, moderate 14–19, severe 20–21). When participants were divided into two groups, with depression scores below or above 10 points, the only differences in life support preferences were found in Case 5 (constant pain) in the immediate evaluation, $\chi^2(1) = 7.71$, $p = .005$, $w = 0.65$. Participants with lower levels of depression more often chose to be fed with a feeding tube (66.66%) than not (8.33%), while participants with higher levels of depression were more evenly divided between...
choosing a feeding tube (11.11%) and being left untreated (13.88%).

There were no significant differences in recall, but in recognition the participants with lower depression levels had more false alarms immediately (.57, .35), \( t(34) = 2.64, p = .012, d = 1.02 \), and at one month (.62, .44), \( t(34) = 2.35, p = .025, d = 0.90 \), than the participants with higher depression scores. The response criterion in the immediate evaluation was also more lenient (–32, .12), \( t(34) = 2.32, p = .027, d = 0.89 \), among participants with lower depression levels. As for gender, there were no significant differences between men and women in either death anxiety or level of depression.

There was no correlation between the scores from the Death Anxiety Scale and Beck Depression Inventory–II; they were not significant when analyzed separately as covariables, nor did they interact with any factor in the recall, recognition or confidence analyses. No significant differences in the Death Anxiety Scale and Beck Depression Inventory–II scores were observed based on whether participants had experienced a serious illness or death of someone close to them or suffered a serious illness or accident themselves.

Discussion

Unlike previous studies about memories for end-of-life decisions, the primary objective of this experiment was to study the memory for the content of the LSPQ, an instrument used to measure end-of-life decisions. The questionnaire includes six very distinct clinical cases (coma, vegetative state, cancer, Alzheimer’s disease, constant pain and kidney failure) that present medical scenarios with serious pathologies, symptoms, diagnoses and treatments. The questionnaire contains emotional material and requires sincere and personal responses. These characteristics make it ideal for examining quantitative and qualitative aspects of the memory for complex clinical situations and treatment preferences. In this study, we analyzed the amount of recalled information, accuracy and false memories in recall and recognition tests. We also measured aspects of metamemory including response confidence and response bias. Lastly, we studied whether variables, including gender, participants’ medical background, fear of death and anxiety have an impact on the memory or life support decisions in the LSPQ.

Recall of the medical scenarios in LSPQ

The participants correctly recalled 50.70% of the contents in the LSPQ in the immediate evaluation. One month later, correct recall came to only 8.20%. Errors increased from the immediate to the one-month evaluation, albeit this effect was smaller than the forgetting of information. The participants had an average of three errors immediately and just over four one month later. Even so, in the one-month recall, the participants gave more than twice as many incorrect units as correct ones.

The poor performance in remembering health scenarios in this experiment is consistent with data on memory for medical information. Forgetting information, inaccuracies and errors are observed in the recall of information that doctors provide their patients (Godwin, 2000; Kessels, 2003) and on the ability of patients to recall and report their medical history (Cohen & Java, 1995). For example, Cohen and Java instructed participants to keep a three-month diary, recording aspects related to health (e.g., symptoms, medication) and rating them on different scales (e.g., seriousness of symptoms). Their participants only remembered 47% of the health-related information, and three months later, performance dropped to 39%. The act of recording, taking notes and assessing the events seems like it would ensure deep and elaborate processing of the information and that the act of processing would improve the recall of medical information, but this was not the case (Onion & Slade, 1995). Therefore, with no elaborate processing to enhance memory for medical content, even poorer performance was expected.

The dramatic loss of information from immediate to one-month recall for the LSPQ could be due in part to the emotionally negative character of contents related to illness, end-of-life decisions and death. We know that details associated with simulations of future negative events are more difficult to remember after one day than details associated with simulations of positive or neutral future events (Ley, 1979). The poor long-term performance may also reflect the influence of the fading affect bias. The intensity of the affect associated with negative autobiographical memories fades faster than the emotion associated with positive memories (Walker & Skowronski, 2009). Perhaps rapid fading of negative affect over time rendered details associated with negative simulations more difficult to recall than those associated with positive or neutral simulations (Szpunar et al., 2012). Moreover, the lack of congruence between the negative emotional state at initial recall prompted by the contents in the questionnaire and the weaker negative state when responding to the memory tests after a month may have impaired long-term recall due to the fading affect (Godwin, 2000). Poor recall of medical information may also have been accentuated with the use of terms that are difficult to understand and remember (Kessels, 2003).

But what did the participants think about their limited one-month recall? Some of them stated that although they tried to remember the contents, they could not recall specific information and were only able to remember general aspects of the cases they were presented with.
These observations are in line with other results that show that the representation of stressful memories or negative events are less detailed or specific (Brown et al., 2014) and are associated with weaker feelings of re-experiencing or pre-experiencing than positive events (D’Argembeau & van der Linden, 2004). The ideas put forward by the participants also support the notion that people may be motivated to stop the retrieval of memories that generate fear, concern or sadness, a blockage that shows the defensive capacity of the self to prevent thoughts that break down personal balance (Conway et al., 2004). Consistent with this idea, inhibitory processes can modulate the accessibility of negative thoughts and episodes and may help keep negative memories from coming to mind (García-Bajos et al., 2017).

The increase in errors from the immediate to the one-month evaluation was significant. There are many aspects that could have influenced the upturn in errors with the passing of time. These include a greater fading affect bias over time (Walker & Skowronska, 2009), a weaker feeling of experiencing (D’Argembeau & van der Linden, 2004), difficulty remembering details associated with negative events (Szpunar et al., 2012), less episodic specificity for stressful memories (Brown et al., 2014) and source errors (Johnson, 1997). An interesting aspect when analyzing the nature of the errors is that they were associated with preconceived ideas or prior knowledge. The findings from the present study show that participants integrate new information from the LSPQ with their own previous knowledge about illnesses (e.g., erroneously remember weight loss in the case of cancer) and experience source confusion between illnesses that have the same symptoms. Both types of errors show the reconstructive processes of memory (Schacter, 2012) and the impact of knowledge schemas on memory (Migueles & García-Bajos, 2012). In summary, time reinforces constructive memory processes.

In the present experiment, when taking into account whether the participants had experienced an illness or death of someone close to them or had personally suffered a serious illness or accident, the only significant differences were between whether or not they had personally suffered a serious illness. The personal implication is important. It has been observed that people who have had real-life complications after an operation remembered less information about the possible complications of the surgery (Godwin, 2000). Similarly, in this study immediate recall was better among participants who had not suffered a serious illness. At one month, there were no differences between whether they had suffered a serious illness or not, but we should keep in mind that correct recall at the one-month evaluation was quite poor, under 10% in both cases. As for incorrect recall, the people who had suffered a serious illness had more errors both immediately and at one month. Therefore, globally speaking, correct recall and errors show that personal experiences negatively affect recall.

**Recognition and confidence of the medical scenarios**

The recall test allows us to examine how emotional information is remembered and to analyze the nature of errors people make spontaneously; the recognition test allows us to see if participants discriminate between true and false information, i.e., whether they accept information that is not part of the LSPQ, but which is coherent or consistent with the given situations. A recognition test is easier than a recall test and requires less cognitive effort (Craik & McDowd, 1987). Studies show that patients are better at recognizing than remembering medical facts. For example, Cohen and Java (1995) observed that participants only remembered 47% of items about health events, but in a recognition test they were able to recover 29% of the facts forgotten in recall. Our study also revealed better results for recognition than for recall. Hits were around 60%, however there were no significant differences between hits and false alarms. This would suggest that participants did not distinguish between correct and incorrect information.

Regarding response bias, participants adopted a liberal response criterion in the recognition test immediately and one month later. They showed a tendency to answer *true* in the recognition test, and accept the content presented, whether true or false. The tendency to accept false content may be explained by the deficits in memory for item-specific information (Brown et al., 2014), poor memory for the source (Johnson, 1997) or dependency of gist and the impact of previous knowledge on decision making (Kessels, 2003). Bias of prior knowledge and schema-driven processing for medical information was particularly significant in recognition. The fact that the false alternatives in the recognition test were plausible and consistent with prior knowledge and schemas may have reinforced this response bias. The recognition test required the content of the emotional situation to be differentiated from the content of our knowledge about illness and treatments. The magnitude of the bias can be large for typical contents or contents consistent with our medical experience because they might receive less specific processing than low-typicality contents.

Do false alarms have the same episodic nature for participants as correctly recognized presented information? Response confidence, although a subjective phenomenological measure, can help us gain understanding of this question. Results show that confidence was rated higher for hits than for false alarms. But confidence does not seem to be a good predictor of accuracy because
hits and false alarms, immediately and at one month, were rated with greater confidence than the mean value of 4 on the scale of 1–7. The participants did not distinguish whether the typical and coherent information was present in the questionnaire or not, and for them the information was true because their decisions were backed by high confidence ratings. Considering false alarms, response bias and confidence, the bias of plausibility of the contents and the impact of prior-knowledge driven-processing was outstanding for medical contents.

Life support decisions and stability

Thirty percent of the participants knew about the ADs document, but only one had actually signed one. After being explained what it was, 35 of the 36 participants in this study said they would be willing to sign one.

In this experiment, there were no significant changes in life support preferences from the immediate to the one-month evaluation, and stability exceeded 82%. Previous studies show that when people are presented with hypothetical end-of-life medical scenarios (e.g., advanced Alzheimer’s or incurable cancer) and asked about their life support preferences (e.g., cardiopulmonary resuscitation) at different times (weeks, months or years later), a moderate consistency is observed (Barrio-Cantalejo et al., 2013; Ditto, Hawkins, & Pizarro, 2005). Stability across all the studies is 71% (range: 57% to 89%). Our results, with an evaluation after one month, are consistent with this data.

Regarding acceptance and refusal of treatments, we found three different results. There was refusal in cases 1 and 2, in which the treatments would not alter the vegetative state or coma, and the participants chose to be allowed to die. The participants accepted treatment in cases 3 and 6 – chemotherapy for cancer and a dialysis machine for kidney failure – and most of them chose not to be allowed to die. There was greater discrepancy between acceptance and refusal of treatment in cases 4 and 5, the amputation of a leg for a person with Alzheimer’s and a feeding tube for the patient with chronic pain. Decisions on whether to accept treatment may have been determined by expectations of treatment effectiveness. When the illness is terminal or there is no hope of recovery, participants choose to end the suffering and die with dignity, but when there is a possibility for improvement, however remote, they choose to pursue treatment (Allmark, 2002).

One aspect that may affect the choice of end-of-life support preferences is whether a person has experienced or personally suffered a serious illness or accident, as supported by previous studies (Barrio-Cantalejo et al., 2013). In the present study, the participants who had experienced illness or the death of someone close to them more often chose chemotherapy to treat cancer, and the participants who had not personally suffered a serious illness or accident more often opted for artificial nutrition and hydration for the clinical case of constant pain. In other words, our results suggest that our past experiences influence our future treatment decisions and preferences.

In terms of gender differences, previous studies with older adults have produced varying results. Men preferred life-sustaining treatments more than women, and women indicated a greater desire for a dignified death than men in one study (Bookwala et al., 2001), while in another, no effects of gender were found in older adults’ preferences (Barrio-Cantalejo et al., 2013). In this study, using university students, the women chose chemotherapy treatment for cancer to extend life at a higher rate than the men. This effect may be associated with the more generalized role of women as caregivers and thus greater involved in medical treatments. This experience and knowledge could have led them to consider the medical scenarios and cancer survival in a more balanced fashion than men. For example, the survival rate in Spain for breast cancer, the most commonly diagnosed cancer in women, is over 80% (Chirlaque et al., 2010).

Death anxiety, depression, and life support preferences

The participants were healthy young adults with low levels of death anxiety and depression. When divided into two groups of higher and lower death anxiety, there were no differences in either life support preferences or any measure of memory. As for depression, participants with lower levels had a more lenient response criterion and more false alarms in recognition than participants with higher levels of depression. Regardless, future studies should examine whether these variables play a role in memory performance for medical aspects. In this regard, it would be interesting to use real-life situations to see if patients can remember precise information about their illness when they have negative diagnoses and poor prognoses that activate the notion of impending death and cause depression. A good memory is needed to understand information provided by doctors and clinical personnel, and memory is a prerequisite for good adherence to recommended treatment (Kessels, 2003).

The memory for the medical information in six clinical cases from the LSPQ and for end-of-life treatment preferences is quite weak and fragile because it becomes hazy in less than a month and changes easily. Nevertheless, the stability of end-of-life support preferences in this study was greater than 80%. Recall data show that the participants only correctly remembered half of the contents of the clinical cases, treatments and consequences immediately, and after one month only just over 8%. These data are consistent with the few
studies that have measured memory for emotional information of a medical nature. As pointed out in earlier studies, it appears that people stop the retrieval of memories that generate fear, worry or sadness and demonstrate a resistance to remembering information about illnesses and medical information. In recognition, our participants had problems discriminating between plausible contents present or absent from the LSPQ. The poor recall of medical scenarios and false alarms in recognizing medical information should be brought to the attention of researchers, medical personnel, and lawyers. In short, professionals should understand that patients have difficulty understanding complex medical information, that memory for this type of content is poor and is badly retained in the long-term, and that real information and generic knowledge is easily confused. All these aspects should be considered since they are key to following medical treatments and can prove decisive in making decisions about complex treatments such as those contained in ADs. Developing strategies and protocols that enhance the understanding and recall of medical information in end-of-life support preferences can be relevant aims for future research.

Since this study was limited to young adults, future research about memory and AD choices should be extended to representative samples of middle-aged and older adults to generalize the results to people of legal age to sign the document. In addition, the fact that the participants were university students means that a replication of the results would be useful to detect individual differences related to the educational levels of the general population. Additionally, it is feasible that special medical or psychological characteristics of the participants can contribute to cognitive processing of emotional information, thus influencing compression and memory of the ADs. In this sense, the results of the present study indicate the contribution of prior experience or knowledge about diseases, personal or relatives’ health conditions, and psychological factors such as depression or death anxiety in the results of memory and end-of-life support preferences. Finally, future research on end-of-life support preferences should take into account the differences found in the distribution of acceptances and rejections of treatments specified in the ADs.

References


D’Argembeau A., & van der Linden M. (2004). Phenomenal characteristics associated with projecting oneself back into


