Problem representation is the mental outcome in which an individual constructs, preserves and interprets information of real-world objects and events (McNamara, 1994). Research on the topic of problem representation shows that people with more experience in a domain (such as physics, statistics, or teaching, among others) represent problems in a different way than people with less experience within the same domain. The general finding is that people with more experience have the ability to recognize and represent the deep, inferential and structural characteristics of the information. People with less experience tend to organize problem representations based on explicit, concrete and often irrelevant aspects of a problem such as specific objects mentioned in the problem, situational context, or wording of the question. In other words, people with more experience generate understandings that go beyond what is explicitly presented. In the current study, we were interested in exploring the hypothesis that even people with little experience in a domain have the ability to see the deep structure of problems but don’t because they are distracted by irrelevant surface-level information (Gäschler, Marewski, & Frensch, 2015; Lee & Anderson, 2001). We were interested in investigating whether we could orient people to ignore surface-level features and, as a consequence, get them to concentrate on the structural aspects of the problem.

The ability to represent problems at a deep level is advantageous because it allows individual to accurately understand what is relevant in problems and subsequently recognize or develop appropriate solutions. This type of
problem representation also leads to greater transfer of learned concepts and principles to situations that are conceptually and principally similar but lack shared superficial characteristics (Chi & VanLehn, 2012; Goldstone & Son, 2005; Hogan, Rabinowitz, & Craven, 2003; Kellman et al., 2010; Nokes-Malach, VanLehn, Belenky, Lichtenstein, & Cox, 2013; Novick & Bassock, 2005; Rittle-Johnson, Star, & Durkin, 2012).

Gaschler et al. (2015) group theories on enabling people to perceive what is important in problems into two general classes. The first they label the bottom-up view. In this view, performance improvement is tied to task practice with specific materials. This performance improvement is often a function of the development of an automatic response to specific materials and is tied to those materials. The second they label the top-down view. In this view, performance improvement is often a function of a voluntary decision on the part of the learner. Because of this strategic perspective of the learner, performance improvement would be expected to be more general and not tied to specific materials.

As an example of the bottom-up view in the context of problem representation, researchers have suggested that it is possible to get people to orient toward the deep without specifically providing them with more knowledge about the domain (Guerlain et al., 2004; Kellman, Massey, & Son, 2010). One strategy is to explicitly give students a lot of practice with certain types of problems so that they can more quickly recognize the appropriate schemata for problems. In other words, the goal is to not teach people how to solve the problem by teaching them the rules of the domain; instead, the goal is to teach them what to look for, or pattern recognition. Recently, there has been research showing that this intervention can have a significant beneficial effect on the ability to perceive the structural aspects of problems (Kellman, Massey, Roth, et al., 2010; Lee, Betts, & Anderson, 2015). For example, Schwartz, Chase, Oppezzo, and Chin (2011) used an “invent-with-contrasting-cases” strategy to facilitate the ability to perceive the structure of ratio in density problems. They observed that this technique facilitated perception of the deep structure and led to increased transfer to other problems in different domains that required perceiving the ratio structure.

While the approach of building up specific problem schemata has merit and empirical support, it also has one major limitation. This approach requires you to practice and build up problem schemata for each different type of problem that students need to learn. Each different type of problem is related to its own problem schema. Each different type of problem would require its own set of practice trials.

In the current study, we explore a slightly different, and perhaps a more general, approach—a top-down approach. This approach assumes that students already have the relevant schemata and that the problem is that they need to learn to orient away from the surface-level features to be able to perceive the deep. In previous research, Rabinowitz and colleagues used a triad judgment task to investigate experience-level differences in problem representation (Fialkov, Jackson, & Rabinowitz, 2014; Hogan & Rabinowitz, 2009; Rabinowitz & Hogan, 2008). The triad judgment task requires participants to identify which of two given source problems “goes best” with a target problem.

Within this task, three different types of triads were developed. The first is the most difficult and is called similar surface-similar deep (SS/SD). With this triad type, a target problem or scenario is presented with one problem that is similar in terms of surface-level features and one that is similar in terms of deep-level features. The participants choose which one of these two problems goes best with the target problem. This triad type assesses whether participants pay more attention to surface-level or deep-level features. The second type is called similar surface-unrelated (SS/U). In this triad type, the target problem is presented with one problem that is similar in terms of surface features and one that is unrelated (doesn’t share deep or surface-level features). This triad type is added as a control to make sure that participants do not just ignore problems that look similar. The third type is called similar deep-unrelated (SD/U). With this type of triad, the goal was to assess whether participants can see the deep-level features if the distracting surface-level similarities were removed. If they are not able to perceive the deep features then both problems should look like they are unrelated and they should just guess.

Studies using three different content areas—statistics, teaching and ethics—have shown consistent findings (Fialkov et al., 2014; Hogan & Rabinowitz, 2009; Rabinowitz & Hogan, 2008). First, people with more experience in a domain do significantly better on triad Type 1 (SS/SD) which is what was expected. People with more experience were more likely to focus on deep-level features, whereas people with less experience focus on the surface-level features. Second, there were no significant differences between people with different levels of experience on Triad Type 2 (SS/U). All participants can see the similarity between the surface-level characteristics of the problems. Finally, we found that with Triad Type 3 (SD/U), participants were able (give that they had some experience in the domain) to perceive the deep-level features of the problems if the distracting surface-level features were removed. We interpret this finding as demonstrating a competence/performance distinction or a production deficiency. That is, participants can see the deep-level similarities when distracting surface-level similarities are removed—they have the competence; however, they don’t exhibit this skill when surface-level similarities are obvious.

The study conducted by Fialkov et al. (2014) was interesting because it investigated people’s ability to perceive
ethical issues within scenarios related to the practice of psychology. An example of a SS/SD triad is shown in Figure 1. In this study, as in the other ones, experience-level differences were observed, but even people with no experience (never have taken a course in ethics within psychology) were able to perceive the ethical issues. Thus, with these materials, we used participants who did not have any formal education within that domain.

The materials used in Fialkov et al. (2014) consisted of 30 triads: 10 SS/SD, 10 SS/U and 10 SD/U. The 10 triads of each type were constructed so that they would involve one of the 10 ethical codes involved in psychology that were presented as the American Psychological Association (APA, 2002, 2010) ethical codes and standards. The ethical standards included Multiple Relationships, Advertising and Other Public Statements, Unfair discrimination, Disclosures, Privacy and Confidentiality, Record Keeping and Fees, Informed Consent, Research and Publication, Avoiding Harm and Competence (APA, 2002, 2010). In the Fialkov et al. (2014) study, the 30 triads were randomly presented to each participant.

If we were to conduct a study like that of Schwartz et al. (2011), to develop an appropriate schema, participants would need to be given a lot of practice with a given type of problem. In the context of the ethics domain, this would mean that they would receive a lot of practice with scenarios that involve a specific ethical issue, such as Informed Consent or Avoiding Harm. However, in the current study, each ethical issue is only presented once. They are not being trained to recognize a specific ethical issue; they are being encouraged to look for the deep, underlying presence of a possible, unspecified ethical issue.

With the competence/performance distinction in mind, we sought to assess whether we could create a context that implicitly encouraged people to exhibit their hidden competence or to heighten it. We hypothesized that the order in which participants encountered the triad judgment types would impact the participant’s ability to properly choose the deep scenario presented in an SS/SD triad type. The order of type encountered, or “blocking,” would orient participants toward either deep or superficial characteristics depending on which triad type they encountered first.
For example, if the first 10 problems they encounter are SS/U, they will be accustomed to choosing superficially similar responses when they encounter the SS/SD triad. To test this hypothesis, we blocked the triad tasks as SS/U, SS/SD, SD/U. However, if the first task they encounter consists of 10 SD/U triad types, they will be accustomed to choosing superficially and become better equipped to disregard the surface-level features of the next problem type. To test this hypothesis, we blocked the triad tasks as SD/U, SS/SD, SS/U. These two are examples of the blocking orders used in the current study.

Method

Participants

There were 313 participants in the study. There were 121 male participants (38.66%) and 192 female participants (61.34%). The participants ranged in age from 18 to 72. The average age within all of the experimental groups was in the 30s and age did not correlate with any of the dependent measures. All participants lived in the United States. The participants were recruited through Mechanical Turk and were paid US$0.25 to participate in the study. The materials and the procedure were approved by the Institutional Review Board (IRB) of the University.

Materials

A triad judgment task was used. This task involved the presentation of three ethical scenarios related to the field of psychology that comprised one target scenario and two source scenarios. Participants were asked to read each scenario and determine which of the two source scenarios “goes best” with the target scenario. Three different types of triads were constructed. Comparisons were based upon two features: surface and deep. Surface features were defined by superficial, narrative similarities to the target scenario. Deep features were defined by sharing ethical constructs with the target scenario. In the first type of triad, one scenario shared similar surface features with the target statement, while the other scenario shared similar deep features (Similar Surface/Similar Deep [SS/SD], see Figure 1, Example 1a). In the second type of triad, one scenario shared similar deep features with the target statement, while the other did not share similarities in either surface or deep features (Similar Deep/Unrelated [SD/U], see Figure 1, Example 1b). In the third type of triad, one scenario shared similar surface features with the target statement while the other did not share similarities in either surface or deep features. Triad 3 was considered Surface Features/Unrelated (SS/U, see Figure 1, Example 1c).

Thirty triads were constructed. One scenario was constructed for each of the 10 ethical principles and those scenarios were then paired so that the three types of triads were represented. Seven different blocking orders were created to assess whether blocking triads to triad types would orient participants to either the surface or deep features. Blocking refers to the order in which participants receive the triad type. Seven different blocking orders were created: (1) SS/SD, SD/U, SS/U; (2) SD/U, SS/SD, SS/U; (3) SS/U, SD/U, SS/SD; (4) SS/SD, SS/U, SD/U; (5) SD/U, SS/U, SS/SD; (6) SS/U, SS/SD, SD/U; and (7) Random or no blocking. The Random condition was just added in as a control condition to assess the effects of blocking versus no blocking and is really not relevant to the goals of this study. As we argue later, this condition does not really contribute to the discussion of the effects of the different blocking conditions. Within each of the six blocking conditions, the triads were presented in a random order.

Procedure

Participants completed the experiment using SurveyMonkey. First, informed consent and demographic information were collected from the participants. Then, participants completed the 30 triad judgment tasks. Each triad was presented on a different page and participants were not allowed to go back and look at triads they already completed. Participants were given no feedback regarding their choices.

Results

Scoring

For each triad, the participant responses were scored as 0 or 1, with 1 point awarded to responses that paired “best” with the target problem. For example, in the SS/SD and SD/U triads, participants received 1 point if they picked the “deep” response. For the SS/U triads, participants received 1 point if they picked the “similar surface” response. For each triad, there were 10 questions, so a perfect score would be 30 points, or 10 points per triad type.

The results were analyzed using a multivariate analysis of variance (MANOVA) with the three dependent variables being performance on the three triad types (SD/SS, SD/U, and SS/U) and the fixed factor being Blocking Order. The overall effect of Blocking Order was significant, $F(18, 1842)=3.504$, $p<0.001$, partial $\eta^2=0.031$. Table 1 shows the means and standard deviations for each of the orders. For the SS/SD triads, the effect of Blocking Order was significant, $F(6, 653)=7.71$, $p<0.001$, partial $\eta^2=0.066$. For the SS/U triads, the effect of Blocking Order was also significant, $F(6, 653)=2.437$, $p=0.024$, partial $\eta^2=0.022$. However, the effect of Blocking Order for the SD/U triads was not found to be significant, $F(6, 653)=0.444$, $p=0.85$, partial $\eta^2=0.004$.

In order to make sense of these results regarding the effects of Blocking Order, we simplified the design by first removing the random group which really is not a Blocking Order. We then combined the two groups that started off with the same Blocking (e.g., SS/SD, SD/U, SS/U; and SS/
SD, SS/U, SD/U). In this way, we could easily assess the
effect of the first experience of the blocking order on the
SS/SD trials. This left us with three groups. Given that
there were no significant effects of Blocking Order on the
SD/U trials, we eliminated those triads from further analys-
es. We analyzed the remaining data by conducting a 3
(Blocking Order—starting SD/SS, starting SS/U and start-
ing SD/U between subjects) × 2 (triad type—D/SS, SS/U—
within subjects) analysis of variance (ANOVA). The
means and standard deviations for this analysis are pre-
sented in Table 2. The analysis yielded a significant main
effect of Blocking Order, \( F(2, 570) = 5.77, p = 0.003 \), partial
\( \eta^2 = 0.02 \); a significant effect of Triad Type, \( F(1, 570) = 346.90, p < 0.001 \), partial
\( \eta^2 = 0.378 \); and a signifi-
cant interaction, \( F(2, 570) = 16.96, p < 0.001 \), partial
\( \eta^2 = 0.056 \).

The most interesting finding from the analysis is the
significant interaction between Blocking Order and Triad
Type. The prediction was that participants who were given
the Blocking Order starting with SD/U would be more
likely to orient their attention to the deep structure when
presented the SS/SD triads and they did and would be less
likely to pay attention to surface-level features; these par-
ticipants had lower scores on the SS/U triads than the other
participants.

**Are people reliably paying attention to the
surface or deep characteristics of the material?**

Absolute differences between the different Blocking Order
groups for any of the triad types do not tell whether people
are reliably orienting their attention to either the surface-
or deep-level information. To assess this, the mean for
each of the triad types within each of the Blocking Orders
were compared to 0.5, which is what would be expected if
people were simply guessing. If the mean is significantly
different from 0.5, this would mean that the participants
were reliably choosing either the deep- or surface-level
choices. If the mean is not significantly different from 0.5,
this does not mean the participants were guessing. It only
means that they were not reliably choosing one of the two
(deep or surface) alternatives. The \( t \)-test values, proba-
bility and reliability of feature are presented in Table 1.

As can be seen in Table 1, for the SS/SD triad type, the
two Blocking Orders that start with SD-U were not signifi-
cantly different from 0.5, although both were at or above
0.5. The other Blocking orders were all significantly below
0.5, which indicates that participants in those Blocking
Orders were reliably choosing the scenario that was simi-
lar in terms of surface-level features. For the SD/U triad
type, all the means were also significantly above 0.5 which
means that participants were reliably choosing the sce-
nario that was similar in terms of the deep structure. This
supports the hypothesis that participants could focus on the
deep structure but don’t do so when there is competition
from surface-level features. Finally, all the means for the
SS-U triad type are also significantly above 0.5, which
suggests that participants were reliably choosing the sur-
face-level alternative for that triad type.

**Was performance in the SS/SD condition
consistent across practice trials or did it
improve gradually with practice trials?**

Given that performance on the SS/SD trials was of primary
interest, we were interested in assessing performance
within the blocking set. Once again, we grouped together
the two groups that started with the same problem block
yielding three groups: those starting with SS/SD, SS/U and
SD/U, respectively. The dependent measure being pre-
sented here is performance on the SS/SD triads. Because
the presentation of the individual triads was presented in a
different random order in each of the conditions, we looked
to see whether performance improved within the block by
comparing the average score on the first five trials with the
average score on the final five trials. The results can be
seen in Figure 2. We analyzed the remaining data by con-
ducting a 3 (Blocking Order—starting SD/SS, starting
SS/U and starting SD/U between subjects) × 2 (triad type—
first five trials, final five trials—within subjects) ANOVA.

The results indicated a significant effect of Blocking
Order, \( F(2, 555) = 20.770, p < 0.001 \), \( \eta^2 = 0.07 \) once again,
showing that participants who started with the SD/U prob-
lems performed better on the SD/U problems than the
other two groups. The effect of triad type was not signifi-
cant, \( F(1, 555) = 1.85, p = 0.174 \), \( \eta^2 = 0.003 \), but a signifi-
cant interaction was observed, \( F(2, 555) = 3.584, p = 0.028 \), \( \eta^2 = 0.013 \). This interaction can be accounted by the
increase in performance across trials for the SD/SS and
SS/U groups and a decrease in performance for the SD/U
group. Participants in the SD/SS and SS/U groups were
going better on the SD/SS trials with practice; however,
participants in the SD/U group started out strong and then
probably started to explore alternative hypotheses as to
what we were looking for.

**Discussion**

The purpose of this study was to replicate and extend prior
research that suggested that the inability of people to orient
to the deep structure of problems is not an issue of not hav-
ing a relevant schema, but rather that they are distracted by
the surface-level features; they are exhibiting a production
deficiency. In fact, the data from the current study clearly
replicate that from the previous studies using the triad
judgment task (Fialkov et al., 2014; Hogan & Rabinowitz,
2009; Rabinowitz & Hogan, 2008). All the Blocking Order
groups were able to reliably choose the scenario that was
similar in terms of deep structure over an unrelated sce-
nario (SD-U triads). If participants did not have the ability

to perceive the deep structure, they would have been in a situation where it seemed like they were being asked to choose between two unrelated scenarios. In addition, five out of the seven Blocking Orders reliably choose the scenario that was similar in surface-level features over those that shared the deep structure when there was a choice between the two.

In terms of extending previous research (Fialkov et al., 2014; Hogan & Rabinowitz, 2009; Rabinowitz & Hogan, 2008), we hypothesized that participants who had experience with the SD/U triad type first would be better able to orient toward the deep and ignore the surface-level features. This hypothesis was also supported. These two groups (SD/U, SS/SD and SS/U; and SD/U, SS/U and SS/SD) exhibited higher scores on the SS-SD triads. Also, while they did not reach a level of score to suggest that they reliably chose the similar deep triad in the SS/SD triad type, they certainly were not reliably orienting to the surface-level features. We also noticed that these participants were scoring lower on the SS/U triad type, suggesting that they were learning to ignore the irrelevant surface-level features.

It is important to contrast the current study from other research that addresses how to get people to see the deep. The most important difference is that we are not teaching people relevant schemata in contrast to other approaches (e.g., Schwartz et al., 2011). In the current study, within the domain of perceiving ethical issues within psychology, there were 10 different ethical issues presented and we never presented more than one example of any given ethical issue. Thus, participants never have the opportunity given this type of practice to develop a specific schema for a given type of ethical issue. We assert that through this type of practice, we are getting students to orient their attention to the deep aspects of the scenario and to pay less attention to the surface-level features. Once again, it is interesting to note that participants who started off with the SD-U triads scored lower on the SS-U triads, suggesting that they were paying less attention to the surface-level features. The orientation presented in this research is not in opposition to the research oriented to teaching specific schemata. Both approaches can accomplish the same goal. The difference is that with the teaching specific schema approach, people would need to teach the specific schema.

Table 1. Means and standard deviations for different triad types as a function of blocking orders.

<table>
<thead>
<tr>
<th>Order</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t value different from 0.5</th>
<th>Probability</th>
<th>Reliable feature orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. SS-SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD/U, SS/SD, SS/U</td>
<td>96</td>
<td>0.533</td>
<td>0.196</td>
<td>t(96) = 1.17</td>
<td>0.099</td>
<td>Surface</td>
</tr>
<tr>
<td>SD/U, SS/U, SS/SD</td>
<td>90</td>
<td>0.501</td>
<td>0.185</td>
<td>t(89) = 0.057</td>
<td>0.996</td>
<td>Surface</td>
</tr>
<tr>
<td>SS/SD, SS/U, SD/U</td>
<td>91</td>
<td>0.440</td>
<td>0.226</td>
<td>t(90) = -2.55</td>
<td>0.012</td>
<td>Surface</td>
</tr>
<tr>
<td>SS/U, SS/SD, SD/U</td>
<td>98</td>
<td>0.348</td>
<td>0.367</td>
<td>t(98) = -4.1</td>
<td>&lt;0.001</td>
<td>Surface</td>
</tr>
<tr>
<td>SS/U, SD/SS, SS/SD</td>
<td>99</td>
<td>0.352</td>
<td>0.271</td>
<td>t(97) = -5.4</td>
<td>&lt;0.001</td>
<td>Surface</td>
</tr>
<tr>
<td>Random</td>
<td>87</td>
<td>0.395</td>
<td>0.217</td>
<td>t(86) = -2.57</td>
<td>0.012</td>
<td>Surface</td>
</tr>
</tbody>
</table>

| B. SS-U      |     |        |      |                           |             |                             |
| SD/U, SS/SD, SS/U | 96  | 0.687  | 0.184| t(96) = 10.02 | <0.001 | Surface                     |
| SD/U, SS/U, SS/SD | 90  | 0.690  | 0.227| t(89) = 7.96 | <0.001 | Surface                     |
| SS/SD, SS/U, SD/U | 91  | 0.746  | 0.229| t(90) = 10.27 | <0.001 | Surface                     |
| SS/U, SS/SD, SD/U | 99  | 0.709  | 0.205| t(98) = 10.14 | <0.001 | Surface                     |
| SS/U, SS/SD, SS/U | 98  | 0.768  | 0.231| t(97) = 11.51 | <0.001 | Surface                     |
| SS/U, SD/SS, SS/SD | 99  | 0.765  | 0.172| t(98) = 15.34 | <0.001 | Surface                     |
| Random       | 87  | 0.713  | 0.241| t(86) = 8.25 | <0.001 | Surface                     |

| C. SD-U      |     |        |      |                           |             |                             |
| SD/U, SS/SD, SS/U | 96  | 0.664  | 0.167| t(96) = 9.57 | <0.001 | Deep                       |
| SD/U, SS/U, SS/SD | 90  | 0.666  | 0.206| t(89) = 7.62 | <0.001 | Deep                       |
| SS/SD, SS/U, SD/U | 91  | 0.655  | 0.215| t(90) = 6.89 | <0.001 | Deep                       |
| SS/SD, SD/U, SS/U | 99  | 0.709  | 0.205| t(98) = 8.53 | <0.001 | Deep                       |
| SS/U, SS/SD, SD/U | 98  | 0.654  | 0.226| t(97) = 6.76 | <0.001 | Deep                       |
| SS/U, SD/SS, SS/SD | 99  | 0.680  | 0.181| t(98) = 9.9 | <0.001 | Deep                       |
| Random       | 87  | 0.641  | 0.199| t(86) = -6.64 | <0.001 | Deep                       |

Table 2. Means and standard deviations.

<table>
<thead>
<tr>
<th>Order</th>
<th>SS/SD</th>
<th>SS/U</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD/U</td>
<td>0.518 (0.191)</td>
<td>0.689 (0.205)</td>
</tr>
<tr>
<td>SS/U</td>
<td>0.374 (0.246)</td>
<td>0.766 (0.203)</td>
</tr>
<tr>
<td>SS/SD</td>
<td>0.392 (0.311)</td>
<td>0.727 (0.217)</td>
</tr>
</tbody>
</table>
for each individual problem type. In the current orientation where we are training people to ignore surface-level features and orient to the deep, we might be developing a more generalized skill that people can apply to a variety of contexts without additional instruction.

Acknowledgements
The authors would like to thank Gregory Murphy, Jenny Arevalo, Ariel Blum and Liat Zabludovsky for helpful comments on an earlier draft of this article. We also like to thank Robert Gaschler and an anonymous reviewer for helpful comments.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


Figure 2. The average score as a function of the initial blocking type and first five versus final five trials.


