The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/191432

Please be advised that this information was generated on 2020-01-22 and may be subject to change.
Uncovering Naturalistic Rewards and their Subjective Value in Forensic Psychiatric Patients

Johanna C. Glimmerveen, Inti A. Brazil, B. H. (Erik) Bulten, and Joseph H. R. Maes

ABSTRACT
The use of relevant reinforcers during treatment is essential for successful interventions. This especially applies to forensic psychiatric populations, which are known to be resistant to treatment. However, it is not clear which rewards are of importance for different types of forensic patients. The aim of the present study was to investigate reward preferences in two forensic patient populations. Applying the concept mapping methodology, 34 male incarcerated violent offenders under imposed psychiatric treatment and 41 male forensic outpatients generated, prioritized and categorized 98 and 115 rewards, respectively. Multidimensional scaling and hierarchical cluster analyses resulted in two concept maps with eight (inpatients) and five (outpatients) reward categories. In both maps, one dimension represented the effort required to achieve the rewards. The other dimension represented either the rewards’ independency of the clinical environment (inpatients) or the level of arousal associated with the rewards (outpatients). Both inpatients and outpatients tended to rate high-effort rewards as the most valuable, especially when the rewards involved the clinical environment of the patient or when rewards were associated with lower levels of arousal. The results highlight the importance of considering individual differences in reward preferences in the development of therapeutic interventions.

KEYWORDS
Forensic psychiatry; offenders; reinforcement; reward value; treatment

The use of reinforcements is a key element of interventions aimed at behavioral modification (e.g., Buehler, Patterson, & Furniss, 1966; Kazdin, 2012; Lussier, Heil, Mongeon, Badger, & Higgins, 2006; Petscher, Rey, & Bailey, 2009), in both psychiatric and non-psychiatric forensic settings (e.g., Timmerman & Emmelkamp, 2005; Wodahl, Garland, Culhane, & McCarty, 2011; Wong, Gordon, & Gu, 2007). Finding relevant reinforcers may be of importance for promoting treatment response in forensic patients, because low treatment responsiveness is common in these populations (Hanson & Morton-Bourgon, 2005; Howels & Day, 2007; Ogloff, Wong, & Greenwood, 1990). Given the growing focus on considering individual differences in responsiveness to treatment (Brazil, van Dongen, Maes, Mars, & Baskin-Sommers, 2016; Insel & Cuthbert, 2015), the use of patient-specific reinforcers is a key step towards developing individualized mental health care and may have positive contributions to the patient’s motivation for treatment.

However, to be effective, reinforcement must be tailored to the needs, preferences, and values of each patient. This is in line with the responsivity principle of the Risk-Need-Responsivity model (RNR; Bonta & Andrews, 2007), which is one of the most prominent theories regarding forensic interventions. The RNR model highlights three principles that promote treatment success: the Risk Principle focuses on the level of risk of the individual offender (i.e., higher-risk offenders will benefit more from more intensive treatment), the need principle highlights the importance of criminogenic needs (i.e., only significant criminogenic factors should be targeted in interventions), and the responsivity principle describes how treatment should be provided (i.e., the intervention should be matched to offender characteristics such as motivational level, learning style, and (inter)personal circumstances).

The RNR model has been criticized for over-emphasizing risk and criminogenic factors, at the expense of attention to individual needs and values (Ward, Melser, & Yates, 2007), which led to the development of the
Good Lives Model as an alternative or complementary approach (GLM; e.g., Ward, 2002; Ward & Gannon, 2006). The GLM states that relapse in risky behavior can be ameliorated by incorporating and emphasizing positive factors during treatment, such as the use of positive reinforcement. The use of positive reinforcement strongly predicts treatment effectiveness within various clinical disciplines (Marshall et al., 2002, 2003).

Importantly, a critical consideration for all forms of reinforcement-based (treatment) programs is that for something to be experienced as rewarding, several requirements must be met: (a) it must be attractive (affective aspect), (b) it must motivate the individual for action (motivational aspect), and (c) the thought of obtaining the reward must lead to an expectation of an enhanced positive state (cognitive aspect; Berridge & Robinson, 2003). This highlights the multi-faceted nature of the experience of reward as well as the need to consider how individual preferences may differ according to these aspects.

There are also more general psychological processes that influence valuation of reward. For instance, the degree of attractiveness of an expected reward is positively related to the amount of effort needed to obtain the reward. This tendency to give greater value to rewards that are harder to obtain has been referred to as effort justification (Aronson & Mills, 1959). It has also been accounted for by the within-trial contrast model (Alessandri, Darcheville, Delevoye-Turrell, & Zentall, 2008; Klein, Bhatt, & Zentall, 2005), which assumes that the value of a reward is dependent on the value of the event that preceded it. In other words, the contrast between a relatively aversive event (such as the exertion of great effort) and a relatively pleasant event (e.g., the reward that follows it) inflates the subjective value of the reward.

Another example is that the temporal distance to the delivery of an expected reward is negatively related to the subjective value of that reward, a phenomenon called delay discounting or temporal discounting (Kirby & Maraković, 1996; Logue, 1988). In other words, individuals generally prefer a smaller immediate reward over a larger delayed reward, but the rate at which the value of the delayed reward decreases differs across individuals. More rapid discounting of rewards is associated with impulsivity-related dysfunctional behavior (Reynolds, 2006), such as substance abuse (Bickel & Marsch, 2001), pathological gambling (Alessi & Petry, 2003), and overeating (Weller, Cook, Avsar, & Cox, 2008). Moreover, rapid discounting is related to antisocial behavior in both children (Barkley, Edwards, Laneri, Fletcher, & Metevia, 2001) and adults (Petry, 2002). It is important to take these discounting effects into account in reinforcement-based therapy, especially in (clinical) populations known to have high discounting rates (e.g., Lussier et al., 2006).

Error justification and delay discounting are examples of how subjectivity affects motivation and reward valuation, depending on individual characteristics and the context in which a reward can be obtained. As such, it should not only be considered how important the use of reward is, but also how reward is experienced, which, in turn, extends to clinical practice. A prominent example of a successful therapeutic approach that incorporates reward and subjective valuation is the Community Reinforcement Approach (CRA), a treatment method for substance use disorders (Hunt & Azrin, 1973). This approach assumes that the use of naturalistic reinforcers with high ecological validity, such as enjoyment of a new hobby, new employment opportunities, or other pleasant activities, promote the development of alternative behavioral styles that are expected to ultimately become more rewarding relative to disruptive tendencies (i.e., substance abusing). One of the main objectives of CRA is to increase the number of healthy social, vocational, and recreational activities that are experienced as rewarding by determining relevant reinforcers for each individual.

Studies on the effectiveness of CRA and its derivatives have yielded positive results in non-residential treatment settings (for a review, see Meyers, Roozen, & Smith, 2011). Adaptations of CRA employ contingency management by rewarding desirable behaviors by, e.g., giving vouchers for drug abstinence (e.g., Secades-Villa et al., 2013). These vouchers are explicit, material reinforcers, and have specific monetary values. Contingency management programs are also used in correctional settings. Points (or vouchers) earned for showing good behavior can be cashed-in to receive material goods or to participate in activities. When such programs are individualized, this may help to decrease both misconduct and reported offender complaints in the prison setting (French & Gendreau, 2006; Gendreau, Listwan, Kuhns, & Exum, 2014; Webb, 2003)

In forensic populations, deficiencies in reward and punishment responsivity have been associated with chronic adult offending (e.g., Buckholtz et al., 2010; Glenn & Yang, 2012), which makes it even more challenging to find well-tailored individualized reinforcers that can be used during treatment, especially in populations of psychiatrically ill offenders. A portion of forensic inpatients has not been continuously engaged in society for a long time, as they have spent time in prison and under imposed forensic psychiatric care. Consequently, they may have developed different needs and thus may experience different stimuli and activities as rewarding compared with individuals who are active members of
society in the outside world. Importantly, the desirable and adaptive behavioral tendencies acquired through treatment must be generalized to life outside of incarceration. However, for forensic inpatients, behavioral change should first be established within the boundaries of the institution. It is therefore essential for in- and outpatients to identify and use those reinforcers that are experienced as being the most rewarding, but are also sufficiently meaningful in their current lives.

Despite the importance of incorporating reinforcers with high subjective value to promote treatment success, it still remains to be assessed which stimuli and activities are generally experienced as sufficiently rewarding by forensic in- and outpatients. To this end, the aims of the present study were to: (a) identify categories of naturalistic rewards that are considered relevant by forensic inpatients and outpatients; (b) investigate how these rewards are valued by these groups of patients; and (c) identify the dimensions across which these rewards can vary.

Methods
Concept mapping
An approach known as concept mapping (e.g., Jackson & Trochim, 2002; Trochim, 1989) was applied to answer the central questions of this study. In general, concept mapping consists of a standardized research protocol that allows the exploration and quantification of how different types of items are clustered into coherent sets and how these sets are mapped on to a higher-order target concept (determined a priori by the experimenter). Ultimately, this approach yields a multidimensional graphical map depicting clusters of items and their interrelationships, and each cluster describes a different aspect of the target concept. It involves collecting qualitative data obtained in processes such as question-driven item generation and unstructured sorting, which in turn are analyzed quantitatively by multidimensional scaling techniques. The concept mapping process used in the present study comprised six stages based on the procedure described by Trochim (1989): (1) participant selection, (2) item generation, (3) item rating, (4) item sorting, (5) statistical analyses, and (6) interpretation.

Stage 1: Participant selection
In order to obtain representative samples of both in- and outpatients, inclusion depended exclusively on the judgment of the head therapists regarding the current mental and emotional stability of the patient. Some of the most prevalent reasons for head therapists advising against recruiting a patient at that particular time were (a) severe depressive symptoms, (b) diminished or disturbed contact with reality (e.g., psychosis), or (c) an intellectual level that was considered too low for constructive contributions to the tasks. Information about the clinical status and index offences of the patients was obtained from their patient files. Sample characteristics, including the most prevalent diagnoses and index offence categories in the patient samples, can be found in Table 1. Psychopathological diagnoses were defined by the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 2000).

JG conducted the recruitment and consent procedures, with the assistance of interns and the clinicians in the respective treatment facilities. All patients were approached for participation only after consulting and having obtained permission from their head therapists. All participants received written and oral information about the study, a financial compensation, and gave written informed consent. Potential participants were allowed a period of at least two weeks to consider and discuss their participation before signing the consent form. The protocol was approved by the local academic ethics committee.

Inpatient sample
Thirty-four male inpatients (item generation: N = 11; item rating: N = 34; item sorting: N = 33), with ages ranging from 21–64 years (M = 43.00, SD = 10.40), and 76% having a Dutch cultural background, were selected from the population of a high security forensic psychiatric institute in The Netherlands.¹ All individuals constituting the inpatient population have committed serious criminal offences in connection with having a DSM-IV axis-I and/or axis-II disorder. Placement in the institute falls under a measure known as Ter Beschikking Stelling (TBS), which is a court-ordered intensive inpatient treatment measure on behalf of the state. TBS can be imposed when the following conditions are met: (a) an offender suffered from a mental disorder at the time of the offence, (b) there is a risk of recidivism due to this disorder, and (c) the offence is punishable by a custodial sentence of at least four years. The inpatient data were collected in 2013 (item generation: N = 11; item rating: N = 15; item sorting: N = 14) and 2016 (item rating and sorting: N = 19).

Outpatient sample
The outpatient sample consisted of 41 male patients (item generation: N = 13; item rating and sorting: N = 31), from three affiliated Dutch forensic outpatient

¹Forensic Psychiatric Centre Pompestichting, Nijmegen, The Netherlands.
treatment centers\textsuperscript{2}, with ages ranging from 20–67 years ($M = 40.02, SD = 12.71$) and 83% having a Dutch cultural background. Individuals in the outpatient sample were engaged in treatment programs focused on problems of aggression regulation or sexual misconduct. Treatment of these patients is voluntary or has been ordered by the court (e.g., as part of probation). The collection of outpatient data took place in 2015.

### Stage 2: Item generation

Items were generated in group sessions that lasted approximately 90 min, focusing on a central question that incorporated the three central elements of reward as outlined by Berridge and Robinson (2003): “What do you regard attractive (affective aspect) to an extent that you would be willing to take effort to achieve it (motivational aspect), because you expect it would bring you a pleasant feeling (cognitive aspect)?” These three elements served as criteria that all items were required to meet in order to be included. In addition, this formulation allowed for the identification of both material and immaterial rewards.

Participants were encouraged to generate as many items as possible, which were directly displayed on a big screen. Both the central question and the generated items remained visible throughout the entire session. When participants came up with items that did not meet all three elements of the central question, they were stimulated to restate the item or to think of related concepts. The same was done when items were considered too vague, too broad, too specific, or when items were formulated negatively (e.g., “Not having (…)”). Highly similar items within one of the two patient samples were removed or merged. When a group of items was considered to consist of examples of a higher order concept, the exemplary items were merged into

---

\textsuperscript{2}Kairos forensic outpatient treatment centers. Locations: Arnhem, Den Bosch, and Nijmegen, The Netherlands.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Inpatients</th>
<th>Outpatients</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>34</td>
<td>41</td>
</tr>
<tr>
<td>Average number of years (SD)</td>
<td>41.97 (9.98)</td>
<td>40.12 (12.71)</td>
</tr>
<tr>
<td>Range</td>
<td>21–64</td>
<td>20–67</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both parents Dutch nationality</td>
<td>76%</td>
<td>83%</td>
</tr>
<tr>
<td>At least one parent Asian, African or South American</td>
<td>24%</td>
<td>17%</td>
</tr>
<tr>
<td>Type of offence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual offence</td>
<td>65%</td>
<td>12%</td>
</tr>
<tr>
<td>Age of victim $\geq$ 16 years</td>
<td>47%</td>
<td>10%</td>
</tr>
<tr>
<td>Age of victim $\geq$ 16 years</td>
<td>18%</td>
<td>2%</td>
</tr>
<tr>
<td>Violent offence (no sexual or material motive)</td>
<td>26%</td>
<td>34%</td>
</tr>
<tr>
<td>Miscellaneous (incl. burglary, theft, stalking, arson)</td>
<td>9%</td>
<td>7%</td>
</tr>
<tr>
<td>No known offence</td>
<td>0% (N/A)</td>
<td>46%</td>
</tr>
<tr>
<td>Legal pressure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Court mandated</td>
<td>100%</td>
<td>41%</td>
</tr>
<tr>
<td>Voluntary</td>
<td>0% (N/A)</td>
<td>54%</td>
</tr>
<tr>
<td>Not yet sentenced</td>
<td>0% (N/A)</td>
<td>5%</td>
</tr>
<tr>
<td>Time in treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\leq$ 3 months</td>
<td>0%</td>
<td>29%</td>
</tr>
<tr>
<td>4-9 months</td>
<td>0%</td>
<td>34%</td>
</tr>
<tr>
<td>10-18 months</td>
<td>6%</td>
<td>20%</td>
</tr>
<tr>
<td>$&gt; 18$ months</td>
<td>94%</td>
<td>17%</td>
</tr>
<tr>
<td>Psychopathology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axis I Depressive disorder</td>
<td>0%</td>
<td>22%</td>
</tr>
<tr>
<td>Attention deficit (hyperactivity) disorder</td>
<td>9%</td>
<td>24%</td>
</tr>
<tr>
<td>Pervasive developmental disorder</td>
<td>18%</td>
<td>5%</td>
</tr>
<tr>
<td>Substance abuse disorder</td>
<td>50%</td>
<td>46%</td>
</tr>
<tr>
<td>Intermittent explosive disorder</td>
<td>0%</td>
<td>37%</td>
</tr>
<tr>
<td>Pedophilia</td>
<td>47%</td>
<td>10%</td>
</tr>
<tr>
<td>Other axis I disorder</td>
<td>24%</td>
<td>34%</td>
</tr>
<tr>
<td>No axis I disorder</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td>Axis II Cluster A personality disorder</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>Cluster B personality disorder</td>
<td>35%</td>
<td>32%</td>
</tr>
<tr>
<td>Cluster C personality disorder</td>
<td>0%</td>
<td>10%</td>
</tr>
<tr>
<td>Personality disorder NOS</td>
<td>56%</td>
<td>41%</td>
</tr>
<tr>
<td>With characteristics of cluster A / B / C</td>
<td>3% / 29% / 21%</td>
<td>2% / 27% / 12%</td>
</tr>
<tr>
<td>Without specified characteristics in record</td>
<td>21%</td>
<td>5%</td>
</tr>
<tr>
<td>No personality disorder</td>
<td>6%</td>
<td>17%</td>
</tr>
</tbody>
</table>

---

\textsuperscript{2}Kairos forensic outpatient treatment centers. Locations: Arnhem, Den Bosch, and Nijmegen, The Netherlands.
one item that was named according to this higher order concept. For instance, the items *Playing soccer*, *Playing tennis*, and *Playing volleyball* were merged into the higher order item *Playing ball games*. A stopping rule was applied after more than 70% of items generated in a single session overlapped the total pool of items generated in previous sessions.

**Inpatient sample**

Two item generation sessions were organized for the inpatient sample. Six patients participated in the first session and five other patients took part in the second session. In the first session, a total of 108 items were generated, of which 54 items (50%) remained in the final list after selection and merging of items by the authors; 150 individual items were generated in the second session, of which 44 items (29%) remained. The final list consisted of 98 individual items.

**Outpatient sample**

Item generation by the outpatient sample occurred in four sessions, in groups of two (sessions 1, 2, and 4) or seven patients (session 3). In the first session, 55 items were generated, of which 42 items (76%) remained in the final list after selection and merging of items by the authors; 72 individual items were generated in the second session, of which 36 items (50%) remained; 42 items were generated in the third session, of which 17 items (40%) remained; and 74 items were generated in the fourth session, of which 20 items (27%) remained. The final list consisted of 115 individual items.

**Stage 3: Item ratings**

All items were numbered and printed on plastic cards. During individual appointments, participants were given a sheet of paper containing five sections numbered 1 (labeled *least important*) to 5 (labeled *most important*). It was explained to participants that their personal attitudes towards the items were of interest and they were then instructed to distribute all rewarding concepts evenly over the five sections. For the inpatient sample, this meant that the 98 cards were to be divided in such a way that two sections would contain 19 cards and three sections would contain 20 cards. Outpatients were instructed to put 23 cards in each section, as their item set contained 115 cards.

**Stage 4: Sorting task**

The sorting task was performed immediately after the rating task. Participants were instructed to group the cards based on similar themes or content in different piles. It was explicitly stated that the sorting should be done in a way that made sense to the participant, and that at least two piles should be created. In addition, the participants gave each pile a label that covered the connection between the items. The sorting data of one inpatient were removed as this participant combined all his categories into one pile at the end of the task and refused to reestablish his previously formed categories. However, the item ratings of this participant (obtained in stage 3) were preserved. Therefore, in the inpatient sample, \( N_{item\ sorting} = 33 \), whereas \( N_{item\ rating} = 34 \).

**Stage 5 and stage 6: Analyses and interpretation**

In order to analyze the sorting data, a binary similarity matrix was created, in which the rows and columns represented the individual reward items (inpatient sample: 98 × 98 matrix; outpatient sample: 115 × 115 matrix). In this matrix, cell values represented the number of participants that placed a pair of items in the same category. Thus, higher scores reflected higher similarities between items. The similarity matrices served as input for nonmetric multidimensional scaling analyses (MDS, PROXSCAL) using SPSS (IBM Corp, 2013), in order to translate the similarity of items into coordinates in a two-dimensional space. For each sample, these coordinates were plotted in a point map, with smaller distances between items on the map reflecting higher similarities.

When evaluating the congruence between the raw data and the final configurations in our study, we deviated from more traditional practices regarding the judgment of MDS solutions in two ways. First, the goodness of fit of the final model is reflected by the stress index, ranging from 0 (perfect fit) to 1 (random configuration). The average stress value of .28 reported in concept mapping studies (Rosas & Kane, 2012) is higher than recommended in the literature on MDS (e.g., Kruskal, 1964), which states that configurations with stress values > .20 are to be considered unreliable (for detailed explanations, see Trochim (1993) and Kane and Trochim (2007)). Importantly, this threshold was established based on simulations and experimental data, which differ fundamentally from the type of data generated using the protocol for concept mapping. Therefore, it seems more appropriate to judge the stress value of our model in relation to results of similar studies, rather than using stress value thresholds obtained from very different data collection protocols (Rosas & Kane, 2012).

The second consideration concerns the primary purpose of the MDS configuration in concept mapping studies, which is to display the clustering results visually. Although a better fit of the data might be observed using more than two dimensions, it would be difficult to generate equally parsimonious and interpretable results in three or more dimensions
(Kruskall & Wish, 1978). In addition, Sturrock and Rocha (2000) showed that two-dimensional MDS solutions have less than a 1% probability of having no structure or a random configuration when stress values are below an upper limit of .39. As we expected our stress value to be close to the average stress value of .28 found in concept mapping studies (Rosas & Kane, 2012), we chose to restrict the MDS analyses in the current study to a two-dimensional solution.

Next, hierarchical cluster analysis using Ward’s algorithm was performed on the MDS coordinates in order to group individual items into clusters, each representing a reward category. The cluster analysis was set at a maximum of 20 clusters and a minimum of 3 clusters (Trochim, 1989). On each step in the analysis, the cluster solution was moved to a lower number of clusters (e.g., from 20 to 19 clusters). The within-cluster sum of squared errors, reflected by the agglomeration coefficient, naturally increases with each step in the clustering procedure. Small coefficients indicate fairly homogeneous clusters, whereas large coefficients or a sudden large incremental percentage change in the coefficient indicates heterogeneous clusters (Hair, Anderson, Tatham, & Black, 1998). The decision on the number of clusters was therefore based on the percentage change in the agglomeration coefficient when moving through the different cluster solutions, as well as on interpretability (i.e., whether a grouping still made sense for the items in the conceptualization).

In order to label the dimensions represented by the axes of the concept maps, the authors independently evaluated the distribution of the items in terms of shared reward features. In other words, it was determined which features were common to rewards on one extreme of an axis and discerned them from those at the other extreme of the axis. For each axis, these features were combined into one dimensional concept that correctly represented all items in the concept map. Finally, average patient ratings for items and clusters were calculated.

**Results**

For clarity and readability purposes, the individual item names have been omitted from the two concept maps. However, the complete lists of reward items are provided as supplemental material.

**Inpatient sample**

The MDS procedure performed on our inpatient data resulted in a final stress value of .26 after 22 iterations. Hierarchical cluster analysis was performed on the item coordinates and inspection of the agglomeration coefficients (see Table 2) shows that the first large percentage change occurs when moving from the eight- to the seven-cluster solution after relatively small increases, and a second jump when moving from the four- to the three-cluster solution. Since (1) the first jump indicates that in the seven-cluster solution two dissimilar clusters have been combined, and (2) the eight-cluster solution was judged by the authors to have the most clearly interpretable reward categories, this indicated that the eight-cluster solution is both statistically and conceptually the most appropriate. Figure 1 shows the inpatient concept map resulting from multidimensional scaling and hierarchical cluster analysis for the inpatient sample. The figure shows the position of the eight reward clusters, the cluster names, and average ratings (between parentheses). Dots represent the individual reward items generated by the inpatients. Lines depict the cluster boundaries.

**Table 2. Hierarchical clustering agglomeration coefficients of the inpatient sample.**

<table>
<thead>
<tr>
<th>Number of clusters</th>
<th>Agglomeration coefficient</th>
<th>Differences in coefficient</th>
<th>Percentage change in coefficient in next level</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1.834</td>
<td>0.183</td>
<td>0.09</td>
</tr>
<tr>
<td>14</td>
<td>2.017</td>
<td>0.239</td>
<td>0.11</td>
</tr>
<tr>
<td>13</td>
<td>2.256</td>
<td>0.274</td>
<td>0.11</td>
</tr>
<tr>
<td>12</td>
<td>2.530</td>
<td>0.31</td>
<td>0.11</td>
</tr>
<tr>
<td>11</td>
<td>2.840</td>
<td>0.427</td>
<td>0.13</td>
</tr>
<tr>
<td>10</td>
<td>3.267</td>
<td>0.526</td>
<td>0.14</td>
</tr>
<tr>
<td>9</td>
<td>3.793</td>
<td>0.624</td>
<td>0.14</td>
</tr>
<tr>
<td>8*</td>
<td>4.417*</td>
<td>1.405*</td>
<td>0.24*</td>
</tr>
<tr>
<td>7</td>
<td>5.822</td>
<td>1.467</td>
<td>0.20</td>
</tr>
<tr>
<td>6</td>
<td>7.289</td>
<td>2.029</td>
<td>0.22</td>
</tr>
<tr>
<td>5</td>
<td>9.318</td>
<td>2.264</td>
<td>0.20</td>
</tr>
<tr>
<td>4</td>
<td>11.582</td>
<td>6.595</td>
<td>0.36</td>
</tr>
<tr>
<td>3</td>
<td>28.177</td>
<td>9.911</td>
<td>0.35</td>
</tr>
<tr>
<td>2</td>
<td>45.329</td>
<td>17.241</td>
<td>0.38</td>
</tr>
<tr>
<td>1</td>
<td>45.329</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *indicates an abrupt large percentage change when moving to a lower number of clusters.
map, depicting the eight reward clusters and their average ratings in two-dimensional space. Exemplary items of the clusters are presented in Table 3.

### Outpatient sample

Performing MDS on our outpatient data resulted in a stress value of .27 after 39 iterations. An abrupt change in the size of the agglomeration coefficients resulting from hierarchical cluster analysis (see Table 4) indicated that dissimilar groups of items were merged when moving from the five- to the four-cluster solution. As the five-cluster solution was also judged to consist of conceptually coherent reward item groups, this solution was identified to be the best representation of the current data. Figure 2 depicts the outpatient concept map, showing the five reward clusters and their average ratings in two-dimensional space. Table 5 presents exemplary items of the clusters.

### Table 3. Cluster names and exemplary items of inpatient rewards.

<table>
<thead>
<tr>
<th>Cluster name</th>
<th>Exemplary items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luxury and material rewards</td>
<td>Salary raise, Coffee</td>
</tr>
<tr>
<td>General social recognition</td>
<td>Promises being kept, Helping others</td>
</tr>
<tr>
<td>Ward climate and restrictions</td>
<td>Visits of volunteers, Possibilities to look for a partner outside the clinic</td>
</tr>
<tr>
<td>Active lifestyle</td>
<td>Swimming, Organizing or attending activities in the clinic</td>
</tr>
<tr>
<td>Substances</td>
<td>Cigarettes, Marijuana</td>
</tr>
<tr>
<td>Autonomy</td>
<td>Taking care of pets</td>
</tr>
<tr>
<td>Relaxation</td>
<td>Singing and making music, Watching a movie</td>
</tr>
<tr>
<td>Leave</td>
<td>Going on leave in the evening, Meeting someone from outside during leave</td>
</tr>
</tbody>
</table>

### Table 4. Hierarchical clustering agglomeration coefficients of the outpatient sample.

<table>
<thead>
<tr>
<th>Number of clusters</th>
<th>Agglomeration coefficient</th>
<th>Differences in coefficient</th>
<th>Percentage change in coefficient in next level</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>2.351</td>
<td>0.272</td>
<td>0.10</td>
</tr>
<tr>
<td>14</td>
<td>2.623</td>
<td>0.275</td>
<td>0.09</td>
</tr>
<tr>
<td>13</td>
<td>2.898</td>
<td>0.327</td>
<td>0.10</td>
</tr>
<tr>
<td>12</td>
<td>3.225</td>
<td>0.455</td>
<td>0.12</td>
</tr>
<tr>
<td>11</td>
<td>3.680</td>
<td>0.628</td>
<td>0.15</td>
</tr>
<tr>
<td>10</td>
<td>4.308</td>
<td>0.631</td>
<td>0.13</td>
</tr>
<tr>
<td>9</td>
<td>4.939</td>
<td>0.733</td>
<td>0.13</td>
</tr>
<tr>
<td>8</td>
<td>5.672</td>
<td>0.83</td>
<td>0.13</td>
</tr>
<tr>
<td>7</td>
<td>6.502</td>
<td>0.973</td>
<td>0.13</td>
</tr>
<tr>
<td>6</td>
<td>7.475</td>
<td>1.222</td>
<td>0.14</td>
</tr>
<tr>
<td>5*</td>
<td>8.697*</td>
<td>4.428*</td>
<td>0.34*</td>
</tr>
<tr>
<td>4</td>
<td>13.125</td>
<td>5.045</td>
<td>0.38</td>
</tr>
<tr>
<td>3</td>
<td>18.170</td>
<td>9.542</td>
<td>0.34</td>
</tr>
<tr>
<td>2</td>
<td>27.712</td>
<td>25.16</td>
<td>0.48</td>
</tr>
<tr>
<td>1</td>
<td>52.872</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * indicates an abrupt large percentage change when moving to a lower number of clusters.

### Table 5. Cluster names and exemplary items of outpatient rewards.

<table>
<thead>
<tr>
<th>Cluster name</th>
<th>Exemplary items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintaining intimate relationships</td>
<td>Having contact with my children on a regular basis</td>
</tr>
<tr>
<td>General social recognition</td>
<td>Being appreciated for my efforts</td>
</tr>
<tr>
<td>Future orientation</td>
<td>Having a job, Making a deadline</td>
</tr>
<tr>
<td>Relaxation</td>
<td>Watching a good movie or documentary, Walking the dog</td>
</tr>
<tr>
<td>Experience seeking</td>
<td>Speeded activities (e.g., quad driving), Holidays and traveling</td>
</tr>
</tbody>
</table>

### Discussion

#### Main findings

The present study identified relevant reward categories for forensic in- and outpatients, as well as the dimensions across which these rewards vary. In addition, preference ratings were collected and mean ratings for each reward category were calculated. Cluster analysis resulted in eight reward categories in the inpatient sample and five reward categories in the outpatient sample. A conceptual comparison of the reward categories of the two samples (see Figures 1 and 2) suggested that both groups identified categories of rewards related to social functioning and personal preferences.
efficacy, as well as more tangible rewards such as concrete items and activities. Not surprisingly, these shared categories concern topics that are important common denominators in the lives of most individuals and are central to classic and contemporary models of human motivation (e.g., Kenrick, Griskevicius, Neuberg, & Schaller, 2010; Maslow, 1943; Ryan & Deci, 2000).

The labeling procedure applied on the axes of the two concept maps resulted in one comparable dimension and one dimension differing between the two samples. Both for inpatients and outpatients, the experience of rewards varied according to the amount of effort required to achieve a reward (y-axis, Figures 1 and 2). Low-effort rewards (e.g., material items or relaxing activities) concerned items that are relatively easy to access, as opposed to high-effort rewards that would require more cognitive and emotional effort and larger time investments. Although this label parallels one of the three reward aspects in the instructions during item generation (based on Berridge & Robinson, 2003), the decision to label this dimension effort was reached by comparing the reward items and their positions in the map. For the inpatient group, the other dimension referred to the degree to which the rewards were independent of the clinical environment (x-axis, Figure 1). In the outpatient group, this dimension concerned the level of arousal associated with the rewards (x-axis, Figure 2).

In both samples, patients tended to rate high-effort rewards as the most valuable, especially when the rewards involved the clinical environment of the patient (inpatient sample) or when the rewards were associated with lower levels of arousal (outpatient sample). One interpretation is that the rewards requiring higher effort were more abstract and long-term rewards that are associated with more intrinsic motivation, which in turn has been shown to be reduced by extrinsic, tangible rewards (Deci, Koestner, & Ryan, 1999; Ryan & Deci, 2000).

The change in cluster ratings when moving along the environment dimension of the inpatients’ concept map shows that the inpatient group rates rewards that apply to their current situation and status in the forensic psychiatric clinic as more valuable than rewards relating to their future lives outside. Low-effort and high-effort rewards related to low independence of the clinical environment are more applicable to patients in later treatment stages. One prediction generated from these findings is that the subjective values of the rewards shift as treatment progresses, so that patients who have the end of incarceration in sight show stronger preferences for rewards that are focused on their lives outside. Further research relating treatment stage to reward (cluster) ratings is needed to explore this hypothesis.

The outpatient group gave higher ratings to rewards that involved lower levels of arousal relative to rewards requiring equal effort but were associated with higher levels of arousal. One explanation of this observation could be that high arousal experiences are perceived as stressful, and as such are linked to negative affective states. Most people will try to avoid experiences leading to these states (e.g., Krieglmeier, Deutsch, De Houwer, & De Raedt, 2010; Rinck & Becker, 2007). Thus, the findings suggest that patients showed a tendency to prefer low-arousal rewards because it seems likely that these rewards were intrinsically associated with more positive affective states relative to high-arousal rewards.

In line with this, literature on positive emotions (Fredrickson, 1998, 2001) suggests that effort and arousal are two dimensions of basic positive emotions such as interest, joy, contentment, and love. Interest, characterized as a high-effort and high-arousal emotion (Fredrickson, 1998), is believed to motivate focused attention, receptivity to information, and learning across situations and throughout the life span (Dougherty, Abe, & Izard, 1996). As such, it shows overlap with the outpatients’ cluster future orientation, which captures items related to personal development, career opportunities and goal setting. Joy, described as a low-effort and high-arousal emotion (Fredrickson, 1998), is associated with the need to be playful, creative, and pushing the limits (Fredrickson, 2001), which parallels the characteristics of the rewards pertaining to the cluster experience seeking. Contentment, a low-effort and low-arousal emotion (Fredrickson, 1998), has been described as a state of inner peace that is felt when people feel comfortable, at ease in, or at one with their situation (Fredrickson, 2013; Mitte & Kämpfè, 2008), which applies to the cluster relaxation. In this theoretical framework (Fredrickson, 1998), love was originally described as overlapping the other emotions and as such was associated with variable levels of arousal and effort. However, when love is interpreted as a mixture of positive emotions in relation to other individuals (Mitte & Kämpfè, 2008), this applies to both the high-effort clusters maintaining intimate relationships and general social recognition of which the items represent low and intermediate levels of arousal, respectively. Thus, the pattern of results converges with the predictions made by the positive emotions framework to explain the link between positive affective states, effort, and arousal.

Moreover, the position of the reward items along the effort dimensions appears related to the extent to which the rewards relate to the pursuit of hedonism and eudaimonia (Ryan & Deci, 2001). Hedonism can be described as a state of immediate, momentary pleasure, with an emphasis on physical stimulation, or a state of relaxation,
whereas eudaimonia can be reached through personal growth and development. Stated differently, hedonism refers to a state of happiness, whereas eudaimonia can be explained as a higher state of well-being. Waterman (1993) found that eudaimonia was more associated with being challenged and exerting effort, whereas hedonic enjoyment was more related to being relaxed, away from problems, and happy. Our findings are in agreement with this dichotomy. Rewards related to autonomy, quality of social functioning, personal growth, and development, are in both concept maps on the high-effort level. Relaxing activities, substance use, and other stimulating experiences are more situated on the low-effort level.

However, it could be argued that criminal activity is often focused on immediate gratification of materialistic desires (e.g., Petry, 2002), or relates to an inability to control inappropriate emotional and sexual impulses (Gottfredson & Hirschi, 1990; Pratt & Cullen, 2000). From that perspective, one could expect that more hedonic rewards would hold higher values in forensic populations, as a hedonistic lifestyle focused on short-term immediate rewards has often contributed to the criminal careers of these individuals. In addition, long-term goals relating to the pursuit of self-fulfillment are often lacking in a large portion of offenders (Pratt & Cullen, 2000), especially in offenders with more severe antisocial tendencies and psychopathy (Hare, 2003; Wiebe, 2003). However, our data show that, at least in these samples, individuals in forensic populations still prefer the achievement of personal growth or development over short-term materialistic rewards.

**Clinical implications**

It is important to keep in mind that determining which reinforcers to use during treatment of offenders with mental disorders is a complex undertaking. To illustrate, reinforcers can be classified as implicit (e.g., personal attention of the therapist) or explicit (e.g., vouchers), short-term (i.e., those effectuated during treatment) or long-term (i.e., the positive effects of successful treatment), and these dimensions will always interact and can even be in conflict. The reinforcers used in reinforcement-based treatments in correctional settings most often involve short-term, explicit, low-effort rewards (Gendreau et al., 2014). Contingency management programs that focus on short-term rewards help promote discipline and structure in prison settings (Webb, 2003), and are effective in the treatment of substance dependence in community settings (Secades-Villa et al., 2013). Regarding long-term behavioral change, our findings are in line with the GLM (Ward, 2002; Ward & Gannon, 2006) and suggest that it may be more beneficial to focus on rewards or goals that increase personal growth and social functioning, which will likely provide more opportunities for success in the future.

The potential of this approach is further highlighted by the fact that offenders released from incarceration are often unprepared for life outside resulting in unemployment, housing problems, drug abuse, and family conflict (Travis, Solomon, & Waul, 2001). Similarly, forensic patients in community settings often experience problems in psychosocial and occupational functioning (Feitsma, Popping, & Jansen, 2010; Henrichs, Bogaerts, Sijtsema, & Klerx-van Mierlo, 2014). In our samples, the rewards increasing eudaimonic well-being were rated as the most valuable, which suggests that forensic patients are willing to provide the effort required to achieve these goals. Therefore, individualized reinforcement programs should not only focus on low-effort, short-term rewards, but should be designed to also include higher-effort, long-term goals in order to further improve treatment success in forensic patients. One approach would be to move from low-effort to high-effort rewards in the course of treatment, as it is likely that while progressing through treatment stages, therapy commitment will increase and treatment goals will become more generalizable to everyday life (Willis, Yates, Gannon, & Ward, 2012). Another step toward the development of personalized reward-based interventions would be to determine the links between patient and offence characteristics, stage of treatment, and the identified reward preferences. It can be expected that different types of forensic patients (e.g., aggressive violent offenders and sex offenders) will differ in what they find rewarding, and that these preferences are also influenced by how responsive they have been to other treatment programs.

**Limitations and recommendations**

It is possible that the recruitment and inclusion process in this study has affected the generalizability of our results to forensic patients in general. Although we aimed to cover the complete range of patient characteristics in our samples, we were dependent on the decision of the patients’ head therapists regarding each patient’s vulnerability and capacity to participate in ongoing scientific research in addition to their daily therapeutic routines. Consequently, our data may not reflect the reward preferences of psychotic, severely depressed, or intellectually impaired patients. More research focused specifically on vulnerable patient groups such as these could shed a light on their respective reward preferences. However, our samples covered a wide range of complex psychopathology, often including comorbid personality and
axis-1 disorders, which in our opinion resulted in reasonably heterogeneous groups.

Regarding the item generation process, a limitation of our study is that the lists of rewards that were generated may not have been exhaustive. However, in order to minimize unnecessary strain on participants, a stopping rule was applied when the amount of overlap between items generated in an individual session and the total pool of items across sessions exceeded 70%. So, it could be argued that we did not identify all of the possible items that are experienced as rewarding in these populations. On the other hand, it seems unlikely that the results would have been significantly different given (a) the large amount of overlap in the items generated in the last sessions and (b) the fact that the number of items identified are similarly distributed across each of the clusters. The latter suggests that additional items would probably also fall within one of the identified clusters and that the concept mapping procedure yielded a relatively complete set of clusters.

A related limitation is that there is a possibility that procedural differences during the item generation stages may have affected the generative process in the two populations. Specifically, more item generation sessions were organized in the outpatient group, because it took longer to reach the stopping criterion. One explanation is that most sessions with outpatients included a lower number of participants, which was mostly due to participants not showing up. Also, whereas there are many parallels across the daily lives of inpatients with respect to, for instance, restrictions and structures, the lives of outpatients are considerably more diverse. Therefore, they may have generated more unique instances of rewarding items that showed less overlap with items from previous sessions.

Another potential issue is that although all three aspects of reward (affective, motivational, cognitive) were required to apply to each generated reward item, it is possible that the item rating procedure may have given unintended extra weight to the affective aspect. Moreover, we do not claim that the current results reflect the only possible outcome. It is possible that other samples of offenders will generate other reward items, even when recruited from the same locations. This is inherent to the subjective nature of the concept mapping process. In addition, the labeling of clusters and axes remains a subjective process and it is possible that we missed other interpretations that could have applied equally well or even better. However, we still found similar clusters (resulting from HCA) as well as one similar dimension (i.e., effort) when comparing the inpatient and outpatient data. This suggests that it is likely that a replication study will yield a similar set of clusters and axes, although the specific reward items may differ to a certain degree from those generated in the current study.

Since we did not include a social desirability measure, it is difficult to say to what degree socially desirable responding can account for the observed general preference for rewards related to personal growth over materialistic rewards. However, both inpatients and outpatients were engaged in intensive therapeutic programs in which personal development is essential to reach the behavioral change required to successfully finish therapy. Moreover, especially inpatients are confronted with severely constrained levels of autonomy and, often, a diminished social network. Consequently, these themes play a prominent role in everyday live, providing a further explanation for the relatively high ratings of related rewards related to these topics.

We did not directly determine whether our data violate the metric axioms that are also required for MDS (Tversky & Gati, 1982). Although the sorting task does not allow violations of minimality and symmetry, it is possible that violations of the triangle inequality principle occur when using similarity judgments. However, the clusters and dimensions were similar between the groups, which is an unlikely finding under severe violations of the triangle inequality assumption because the configuration of points in the MDS map would be highly distorted. This can be seen as an indication that triangle inequality did not have a large effect on our results. Nonetheless, it is important to consider the potential impact of triangle inequality in future studies, which may further aid reproducibility. Finally, the study did not include a healthy control group, which makes it difficult to determine the extent to which the identified clusters differ from what is found rewarding in the general community. It could be very useful to uncover differences in reward preferences between healthy individuals and outpatients, as they have access to similar resources and activities as healthy non-patients in everyday life. Moreover, while this may be less informative for the inpatient population due to the restrictions imposed by institutionalization, it would still be interesting to explore common factors in the preferences of forensic patients in general that differ from those of healthy individuals. In the future, identifying how these groups differ could provide a reference point for clinicians when determining the areas that need to be targeted through personalized treatment.

Conclusions

In summary, to our knowledge this is the first study investigating reward preferences in forensic in- and out-patient populations. Using a unique mixed-methods approach, we found that both inpatients and outpatients
tend to rate rewards requiring greater effort (e.g., rewards related to autonomy, quality of social functioning, or personal growth and development) as more valuable than low-effort rewards (e.g., substances, material goods, relaxing activities, or stimulating experiences), especially when the rewards involve the direct environment of the patient or lower levels of arousal. The findings may foster the development of individualized treatment plans that incorporates a patient’s reward preferences. One scenario is that clinicians could use (a subset of) the reward categories that emerged in the current study to aid in developing individualized reward schedules with their patients. For instance, helping their patients to think of examples in each of the reward categories may provide a structured framework to contemplate which elements in the patient’s life are, or could be, rewarding. A next step could be to make the patient rate the items in terms of attractiveness (i.e., subjective reward value) and to determine together what would be needed to obtain these rewards; the feasibility, the conditions to be met, whether the rewards can be obtained on a short term or a longer term, and so on. Although the therapeutic impact of such an approach will need to be examined, the present study offers a first step toward achieving this goal.

Acknowledgements

The authors would like to thank Barbarien Bruens, Loïs Buyné and Mareike Eckert for their work regarding the collection of sorting and rating data.

Funding

IAB was supported by a VENI grant (451–15–014) from the Netherlands Organisation for Scientific Research (NWO).

ORCID

Johanna C. Glimmerveen http://orcid.org/0000-0002-9254-4325
Inti A. Brazil http://orcid.org/0000-0001-5824-0902

References
