Universiteit Leiden
The Netherlands

The effects of reward systems in prison: a systematic review

Elbers, J.M.; Ginneken, E.F.J.C. van; Nieuwbeerta, P.; Boone, M.M.; Palmen, J.M.H.

Contents lists available at ScienceDirect

International Journal of Law, Crime and Justice

journal homepage: www.elsevier.com/locate/ijlcj

The effects of reward systems in prison: A systematic review☆

Jan Maarten Elbers [*], Esther F.J.C. van Ginneken, Paul Nieuwbeerta, Miranda Boone, Hanneke Palmen

Leiden University, Leiden, the Netherlands

A B S T R A C T

Introduction: Using rewards to influence behavior and attitudes of incarcerated individuals is historical, widespread, and becoming increasingly popular. Surprisingly, little is known on the effectiveness of reward systems in prison (RSPs). This paper aims to address this gap. *Method.* A systematic review on RSPs was conducted, twenty one studies met the inclusion criteria. *Results.* Overall, findings on RSP effectiveness were mixed. Three study types were identified. Type-1 studies indicated that experimental RSPs using tokens yield positive short-term behavioral effects in individuals cognitively functioning below average. Type-2 studies indicated that non-experimental RSPs are most successful when targeting high-risk participants and include cognitive program elements. Type-3 (qualitative) studies highlight three potential moderators: a pain of self-government, perceived legitimacy, and using family contact as a reward. *Discussion.* As overall study quality is below average, findings should be interpreted with caution. Results were partially congruent with prior research.

Reward systems are commonly used in prisons worldwide in order to influence the behavior and attitudes of incarcerated individuals. Reward systems in prison (RSPs) apply the technique of providing rewards contingent on display of 'good' behavior and removing rewards following misbehavior (Burdon et al., 2003). RSPs do not use any other type of punishment. Good behavior (e.g. hygiene maintenance, participation in work or activities) in prison has long been rewarded with sentence discounting (Demleitner, 2017; Hyneman, 1926; Weisburd and Chayet, 1989) and incremental degrees of freedom and privileges (e.g. Gillin, 1931; Hamels, 1996; Maconochie, 1859). Today, rewards are used to manage and change behavior in prison (units) in Canada (Serin and Hanby, 2009) and the United States (Michigan Department of Corrections, 2020; Mitchell, 2010; Rawlinson, K. 2019), and in the entire prison systems of England and Wales (Liebling, 2008), Romania (Morar et al., 2019), the Netherlands (Dutch Prison Service DJI, 2013), and Ireland (Irish Prison Service, 2012). Considering tens of thousands of offenders are sentenced to prison in these countries each year, these reward systems potentially impact the behavior and attitudes of many incarcerated individuals.

RSPs can be an attractive management tool for encouraging compliance, for the benefit of order and safety in prison. It is believed that rewarding individuals with, for example, extra opportunities for visitation or recreation, motivates them to behave well (see also the section 'Rationale of Contingency Management Systems' below). Additionally, it sends the message that resources are only spent on individuals who have earned it, which can be considered politically persuasive. Indeed, this take on treating incarcerated individuals is in line with the current trend of responsabilization. This neoliberal governance strategy refers to transposing responsibilities of

rehabilitation and reform from the government to incarcerated individuals (Garland, 1996, 2001; Hannah-Moffat, 2005; Shammas, 2014Sampson et al., 2007).

Surprisingly, very little is known on RSP effecs on incarcerated individuals' behavior and attitudes, since published reviews of the empirical literature on this topic are nearly absent (Burdon et al., 2003; McGuire, 2018). The only exception – known to the authors – concerns a meta-analysis by Gendreau et al. (2014). The authors examined the effectiveness of one type of RSP: token economies. These systems reward good behavior with tokens, which can be exchanged for social, material and active rewards (e.g. extra visits, cigarettes and access to the gym). Their findings suggest that token economies can successfully change behavior of incarcerated individuals in 69% of cases. However, the included studies do not provide insight into the effectiveness of RSPs in which people are not allowed to select their own rewards, nor distinguishes effects between RSPs targeting different behaviors and populations. This could be essential, as prior empirical research on reward effectiveness has shown that being able to select rewards which are attractive to an individual can moderate RSP effectiveness in terms of individual behavioral and attitudinal change (Lee et al., 2010; Premack, 1965; Sulzer-Azaroff and Mayer, 1991). Other variables presumably affecting the effectiveness of systematic rewarding are system monitoring, complexity of target behavior (Kazdin, 1972; Sulzer-Azaroff and Mayer, 1991) and target population; as participants have to possess sufficient (cognitive) capacities to be able to change and regulate their behavior and be responsive to rewards (Andrews and Bonta, 2010; Kazdin, 1982; Schunk and Zimmerman, 2012; Sulzer-Azaroff and Mayer, 1991).

Of course, numerous reviews on the effectiveness of interventions on behavior and attitudes of incarcerated individuals have been conducted (for a non-exhaustive overview, see Lipsey et al., 2007), yet they do not provide insight into RSP effectiveness. Unfortunately, many of these overarching reviews do not differentiate between reward and punishment techniques in discussing intervention effectiveness (Andrews et al., 1990; French and Gendreau, 2006; Garrett, 1985; Keyes, 1996; Koehler et al., 2013; Landenberger and Lipsey, 2005; Lipsey et al., 2007; Morgan and Flora, 2002; Pearson et al., 2002; Wilson et al., 2005). Be that as it may, offender rehabilitation experts pose that rewards are more effective in changing (antisocial) behavior than punishments (Gendreau et al., 2014), in line with empirical research on this topic (Azrin and Holz, 1966; Marlowe and Kirby, 1999). Some of these reviews do not differentiate between criminal justice settings either (e.g. Lipsey et al., 2007; Pearson et al., 2002). However, empirical prison research indicates that institution-level factors, such as social prison climate, quality of staff-prisoner relationships and security level, can mediate rule compliance (Bosma et al., 2020; Huebner, 2003; Pappas and Dent, 2021; Steiner and Woolredge, 2008), suggesting that prior empirical findings, such as among drug court participants and probationers, might not translate well to the unique prison setting. Furthermore, reviews mostly examined effectiveness in terms of recidivism post-release, but not behavioral and/or attitudinal change in prison. Those that do, report inconclusive outcomes (French and Gendreau, 2006) or small reductions in prison misconduct (Keyes, 1996), and a considerable advancement in a range of psychosocial outcomes (e.g. self-esteem) (Morgan and Flora, 2002).

The lack of an overview of empirical knowledge on reward systems in prison and the lack of consensus on its effects can be explained by the fact that this strand of research has primarily focused on non-incarcerated offenders, specifically substance abusers, probationers and parolees. This research has indicated that (removal of) rewards (often money vouchers or prizes; Petry, 2000) can stimulate therapy engagement, reduce substance related offences (Burdon et al., 2003; Petry, 2000) and promote abstinence in substance abusers (Ainscough et al., 2017; Davis et al., 2016; Griffith et al., 2000; Lussier et al., 2006; Prendergast et al., 2005), but not always (e.g. Hall et al., 2009). Effect sizes were found to be moderated by reward volume and immediacy of reward application (Lussier et al., 2006). In probation and parole, rewards are most commonly used to increase parole-meeting attendance and reduce recidivism. However, an extensive meta-analysis has shown their effectiveness on these two outcome measures to be ambiguous (Prendergast et al., 2015). At best, recidivism seems to be delayed by the use of rewards in post-prison supervision (Sloas et al., 2019). At the intersection of prison and community are sentence discounting programs (e.g. good or earned time credit). The available research on this topic suggests that granting or removing good time, has no or even negative effects on both misconduct in prison and recidivism post-release (Emshoff and Davidson, 1987; Johnson and Stageberg, 2014; Steiner and Cain, 2014; 2019). The extent to which these specific findings among non-incarcerated offenders can be transposed to those in prison, is unclear.

Therefore, in this paper we aim to create more insight in the effectiveness of RSPs, by addressing the following research question: What are the effects of reward systems in prison (RSPs) on the behavior and attitudes of incarcerated individuals? To answer this research question, we conduct a systematic review of the international empirical literature on this topic. With this review, we aim to identify the available knowledge on the effectiveness of reward systems in prison, assess its methodological quality, and identify helpful new avenues for empirical research in this field.

An overview of studies on RSP effectiveness is both theoretically and practically relevant. RSPs are assumed to yield positive effects on behavior. Scholars have even hypothesized that encouraging incarcerated individuals to earn rewards with good behavior can promote their self-rehabilitation through reinforcing prosocial behavior, and also by offering motivated individuals extra education and training (Gendreau et al., 2014). In contrast, other scholars argue that rehabilitation in prison can be frustrated by making access to known promotors of behavioral change dependent on their behavior, such as qualitative social relationships in prison (Craig, 2004), access to visitation (Hutton, 2017) and other activities evidenced to promote rehabilitation (Prison Reform Trust, 2014). An overview of the empirical studies on this topic can provide data to assess these hypotheses.

Additionally, the aim of this systematic review is to generate actionable knowledge for policymakers in the criminal justice system and professionals working in prisons. Providing an oversight of studies on RSP effectiveness can aid criminal justice policy decision-making and professionals working in prisons to implement and operate reward systems effectively. In answering our research question, we discuss the target populations, target behaviors, and techniques of the three types of studies on RSPs which emerged from our findings. Before focusing on the effects of reward systems in prison (RSPs) on behavior and attitudes, we briefly discuss the mechanisms behind possible effects of these systems. In other words, we discuss the theoretical framework that is commonly assumed to underpin RSPs.

1. Rationale of contingency management systems

The assumption that systematic rewarding influences the behavior of incarcerated individuals can be traced back to contingency management systems, of which RSPs are one example. Contingency management systems are based on the principles of operant or instrumental conditioning, which underpin classical behavior therapy (Skinner, 1938). Operant conditioning principles are used in many areas of life, such as pet training, classroom management, and work productivity. Its core assumptions are that behavior is determined by anticipation of its consequences and that the cause of behavior lies in environmental stimuli instead of in the actor himself (Ayllon and Milan, 1979). Pleasant consequences of behavior are believed to increase the likelihood of its recurrence and unpleasant consequences its extinction (Azrin and Holz, 1966; Skinner, 1938). According to the principle of extinction, removing reinforcement will generally cause the performance of previously rewarded behavior to fade out (Vurbic and Bouton, 2014). Following this theoretical perspective, rewards may serve as incentives of compliance in prison, up to the point of release (due to removal of reinforcement). The behaviors which contingency management aim to increase, are referred to as target behaviors.

Four types of rewards and punishments can be distinguished: positive and negative reinforcement,[1] and positive and negative
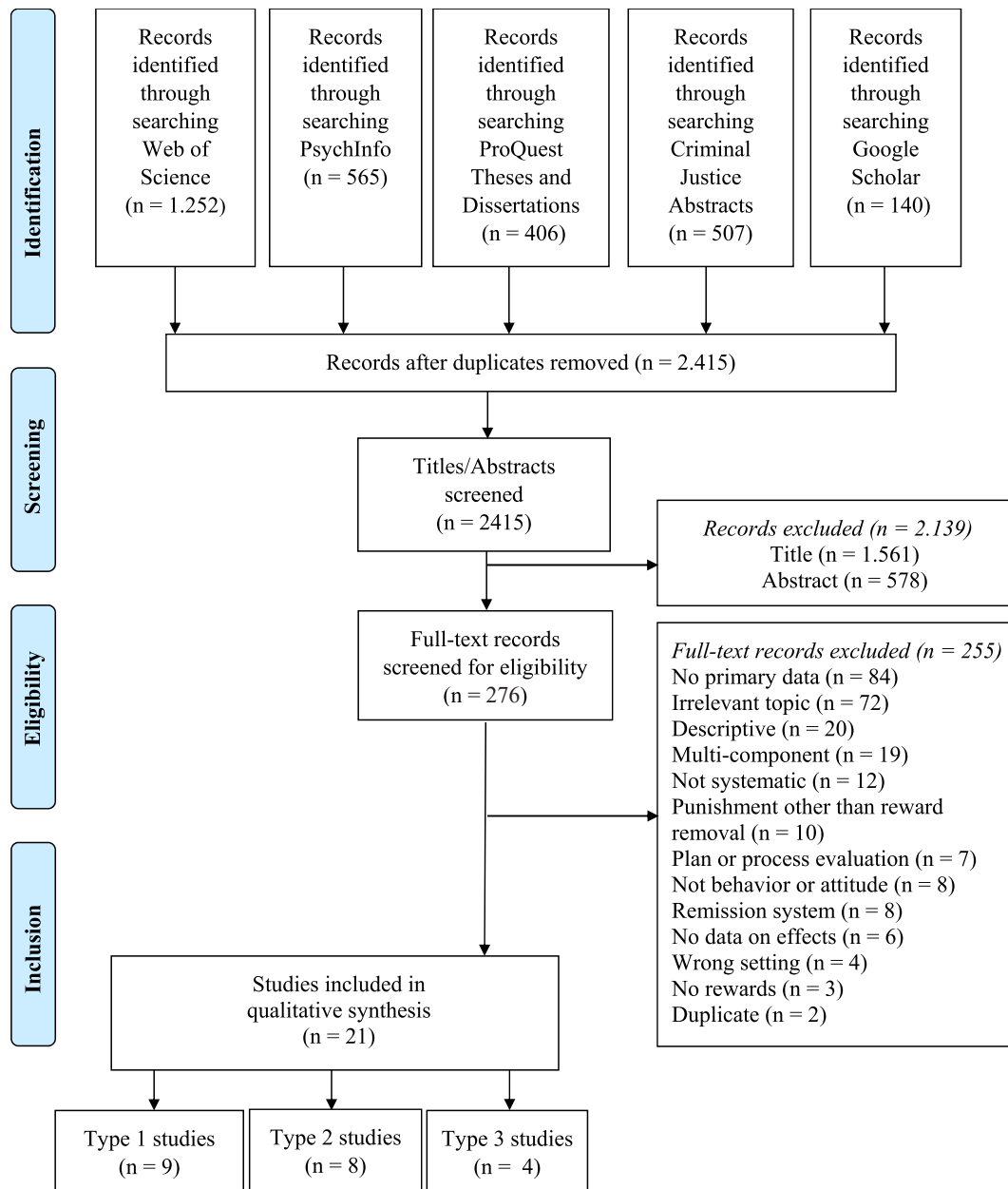

Fig. 1. Prisma flowchart.

punishment. The labels 'positive' and 'negative' are not morally laden, but simply refer to functions. Positive reinforcement translates to adding something positive contingent on performing a desirable behavior (e.g. rewarding hygiene maintenance with extra meals), whereas negative reinforcement refers to removing something unpleasant (e.g. unlocking the cell between activities). The goal of reinforcement is to reinforce the behavior performed. Positive punishment translates to adding an aversive consequence after performing an undesirable behavior (e.g. isolation following an incident), whereas negative punishment refers to removing a reinforcer (e. g. removing visitation privileges following an incident). Negative punishment is also referred to as response cost (Kazdin, 1972). The goal of punishment is to decrease the future occurence of behavior performed. Contingency management systems can use a selection of all four types of rewards and punishments. However, reward systems in prison exclusively use positive reinforcement (reward awarding) and negative punishment (reward removal). Several techniques for reward application can be used, such as the use of tokens or points which can be exchanged for rewards, pre-selected sets of rewards that are provided upon level progression, and the use of contingency contracts in which mutually agreed upon target behaviors and rewards are described, upfront.

RSPs thus rely heavily on extrinsic motivation as a driver of positive behavior, but this may be at the cost of intrinsic motivation to change. Previous research found that external regulation of behavior and extrinsic rewards (e.g., in a school context) may undermine intrinsic motivation (Deci et al., 1999). Given the popularity of RSPs, the lack of a clear evidence base, and potentially adverse and undesirable effects, research on the application and effectiveness of RSPs is timely and necessary.

2. Methodology

The main goal of this study was to assess the effects of RSPs on the behavior and attitudes of incarcerated individuals. Therefore, we conducted a systematic review following these recommended steps: identification, screening, eligibility, and inclusion (Higgins and Green, 2021; Petticrew and Roberts, 2008). The entire screening process is visually captured in the PRISMA flowchart (Fig. 1) (Moher et al., 2009).

2.1. Identification of possibly relevant publications

First, we searched for relevant publications in several academic research databases. Given the topic of this study and our research question, we chose to identify relevant records by searching scientific databases Web of Science (1945-present), PsychInfo (1967-present), Criminal Justice Abstracts (1968-present), ProQuest and Google Scholar. Web of Science (which hosts a collection of databases, e.g. MEDLINE and ScIELO) was selected as it is the world's leading multidisciplinary scientific search tool. PsychInfo complements this database as it has been found to include studies on psychological and psychiatric topics that are not available in other databases, such as MEDLINE (e.g. Brettle and Long, 2001; Stevinson and Lawlor, 2004). To avoid publication bias, Criminal Justice Abstracts (which includes dissertations, government reports, books, and unpublished papers), ProQuest Dissertations and Theses, and Google Scholar were also searched. No temporal search limits were set. The entire search was conducted in March 2021. We conducted the literature search with specific inclusion criteria – and thus search terms – that were based on the central research question.[2] Therefore, publications were only considered relevant if these described a study in which (a) the reward system was applied in a prison or jail, (b) the reward system involved providing and/or removing rewards, (c) the rewards were provided based on an assessment of a defined target behavior, (d) the effects/outcomes discussed were attributed to the RSP and (e) rewards were at least partially enjoyed in prison. Study quality was not an inclusion criterion, as we aimed to map all relevant research on this topic and assess its methodological quality. Additionally, to further minimize publication bias, we snowballed through references of the included studies and important retrieved literature reviews and meta-analyses on contingency management systems in incarcerated populations (e.g. Gendreau et al., 2014; Serin and Hanby, 2009). Snowballing did not reveal additional unique records. These search strategies combined resulted in 2972 records. After removal of 557 duplicates, 2415 unique records were identified.

2.2. Screening title and abstract

Second, each of the 2415 unique publications was screened by title or abstract and selected when relevant. The aim of screening was to exclude publications that did not meet the inclusion criteria. As a result, reward systems in other settings than prisons or jails, such as forensic hospitals, reformatory schools, problem-solving courts, parole and probation, were excluded. Studies on remission systems, good and earned time credit, and early release programs were also excluded. This also applied to RSPs that only use positive punishment, and RSPs introduced as part of a non-contingent multimodal treatment program. Yet, RSPs in which engaging in such programs was a specific target behavior and thus rewarded, were included. The screening of all unique 2415 publications, resulted in 334 potentially relevant studies, of which 79 full-texts were not located. The remaining number of 255 studies is substantial, as we chose to screen full-texts when in slightest doubt of relevance. And doubt there often was, because titles and abstracts often contained terms such as 'behavior modification', 'reinforcement' or 'incentive', whereas those labels did not reveal whether it concerned reward (removal) or other forms of reinforcement and/or punishment. Although labor-intensive, this approach was chosen to circumvent unfortunate exclusion of relevant studies.

2.3. Screening full-texts

Third, the full-texts of the remaining 255 studies were read, and publications were excluded when they did not meet the inclusion criteria. Again, the same inclusion criteria were used as in the first and second step. Note that quantitative studies, mixed-methods and

qualitative studies were included. Inclusion of qualitative studies in systematic reviews is not common in criminology (Ajzenstadt, 2016). However, doing so was relevant to our research question, as qualitative insights could potentially be helpful for understanding effects found in quantitative studies (Seers, 2015) and identifying (undesirable) side effects. The screening of the full-texts of 255 studies resulted in a selection of 21 relevant studies. Two of those studies were reported in one publication (Milan et al., 1979b). Two separate publications reported results from the same sample (Khan 2016, 2020). Most full-texts were excluded for reasons of not containing primary data, full-texts not being available or being irrelevant upon further inquiry (see Fig. 1).

2.4. Synthesis

Finally, all 21 studies that met the inclusion criteria were synthesized, focusing on study outcomes and quality. Synthesis involved a critical appraisal of the results. For this purpose, we extracted data from the studies. For quantitative studies, a data extraction form was prepared with key meta-data (e.g., sample characteristics, setting, reward types, reported effects, etc.), which was analyzed for patterns. The same method was used for qualitative studies, except that the results consisted of themes that were described and discussed by the author. No additional content analysis of these studies was conducted to see if other themes would emerge. When reading the selected 21 studies, a pattern emerged in the sense that three types of studies could be distinguished, and we classified each study into one of these types. The three distinct types of studies on RSPs are: (1) quantitative studies conducted before 1993, (2) quantitative studies conducted after the year 2000 and (3) qualitative studies conducted after the year 2000.[3] After reading all 21 studies carefully, we classified 9 studies as type 1, 8 as type 2 and 4 as type 3. Note that the three types of studies do not only differ according to their publication date or whether these are qualitative or quantitative, but more importantly in context, sample size, dominant research method, research design and technique, as well as immediacy of reward application (see our discussion below and Table 1).

2.5. Quality assessment

The Maryland *Scientific Method Scale* (SMS) (Cook and Campbell, 1979) was used to assess study quality of quantitative studies, which is commonly used in criminology (Farrington et al., 2003; Petticrew and Roberts, 2008). This scale is a five-point scale ranging from 1 (cross-sectional correlations) to 5 (randomised-controlled trials). Six studies were scored as 1, as these cross-sectional studies lacked control groups. We scored the quality of seven studies as 2, among which were ABA-designs, which included before-after comparisons, but lacked counterfactuals and randomized participant selection. Three studies were scored 3, as they included before-after comparisons and control groups, but important unobserved differences were likely to exist between these groups. One quasi-experimental study (before-after comparison, matched (but not randomized) control group, use of control variables) was scored 4. That no randomized-controlled trials were found is in line with previous findings indicating the scarce use of this golden standard to study intervention outcomes in prisons (e.g. Lipsey et al., 2007; Beaudry et al., 2021). Thus, overall study quality of quantitative studies was low to medium in terms of SMS-scores. However, this does not necessarily mean these studies do not contribute to our understanding of reward systems in prison. On the contrary, cross-sectional study outcomes can be valid when the exposure to the intervention can be assumed to be stable over time (Kesmodel, 2018), which can be argued to be the case for these prison-based studies. Although hard to draw robust conclusions on causality based on low-to medium-score studies, they did present valuable information on correlations between systematic rewarding and behavioral or attitudinal outcomes.

For assessing quality of qualitative studies, we used the Critical Appraisal Skills Program (CASP) checklist for qualitative research (Critical Appraisal Skills Programme, 2018). The CASP is a commonly used tool for study quality appraisal, supported by the Cochrane Qualitative and Implementation Methods Group (Long et al., 2020) and contains questions which are included in most study-quality appraisal tools (Munthe-Kaas et al., 2019). This checklist contains ten review questions (e.g. 'was the research design appropriate to address the aims of the research?'), which have to be answered with 'yes', 'can't tell' or 'no'. A common difficulty in the use of CASP is how to distinguish between genuine research flaws and lack of reporting (Long et al., 2020). Fortunately, this issue was easily

Table 1
Characteristics of three types of RSP studies.

Characteristics	Type 1	Type 2	Type 3
Study			
Number of studies	9	8	4
Publication date	Pre or in 1993	Post 2000	Post 2000
Context	US	US, UK, Romania	UK
Sample size	Range = 2 - 56	Range = 10 - 4960	Range = 16-72
Methodology	Quantitative	Quantitative, mixed,	Qualitative
Participation	Part voluntary, part imposed	Imposed	Imposed
System			
Design	Experimental	Non-experimental	Non-experimental
System	One-leveled token economies	One-leveled token economies, multiple level progression-dependent reward systems	Three-level progression-dependent reward systems
Immediacy of rewards	Immediately following target behavior(s)	1–120 days following target behavior(s) and/or attitudes	14–28 days following target behavior(s) and/or attitudes

overcome, as all of the included qualitative studies were part of larger dissertational research projects, all of which provided additional information relevant for assessing individual study quality. Another issue of using CASP is how to arrive at overall quality judgements; determining which or how many no's indicate poor quality (Long et al., 2020). As most questions for most of the included qualitative studies were answered affirmatively, and no's did not relate to any fatal flaws (Dixon-Woods et al., 2004) or essential elements (Carroll et al., 2012a, b) – such as rigor of data collection or appropriateness of using a qualitative methodology – we concluded that all four studies were of good quality.

3. Results

3.1. Characteristics of studies on RSP effectiveness

As discussed above, our systematic literature review resulted in 21 studies, and three types of studies were distinguished (appendix 1 lists all study key characteristics). Nine studies were categorized as type 1: quantitative studies conducted before 1993, eight studies were categorized as type 2: quantitative studies conducted after the year 2000, and four studies were categorized as type 3: qualitative studies conducted after the year 2000.

The three types of studies clearly differed in their study characteristics (see Table 1). Type-1 studies were characterized by (i) exclusively being conducted in the United States, (ii) with relatively small sample sizes, (iii) using quantitative research methods. More importantly, however, all studies concerned experimental token economies, in which rewards were provided immediately following display of good behavior. Six of these studies were conducted in a special experimental unit in Draper correctional center in Elmore Alabama, in the 1960s. Rewards were awarded immediately upon display of target behavior. Four out of nine studies were included in a prior meta-analysis (Gendreau et al., 2014). Apart from the RSP in one study (Ellis, 1993), all systems were developed, implemented and monitored by external researchers. All type-1 studies monitored and tracked behavior of single subjects.

Type-2 studies were characterized by (i) being conducted in various countries (predominantly United States and the United Kingdom), (ii) with relatively large sample sizes, (iii) using quantitative research methods, (iv) to study non-experimental RSPs, (v) which use non-immediate rewarding. In contrast with most type-1 studies, type-2 studies mostly included participants who did not volunteer for system participation. Moreover, two out of eight systems in these studies were developed, implemented and monitored by external researchers. The remaining six studies were conducted by prison administrations. Three out of eight systems in these studies involved points or tokens. Time between display of target behaviors and receiving rewards was usually less immediate than in the type-1 studies, as this period ranged from 1 to 7 days (Feinstein, 2003), 14–28 days (Liebling, 2008, see British Ministry of Justice, 2011), 14–120 days (Reid et al., 2000), 30 days (Sharma and Marino, 2017), 30–60 days (Meyers et al., 2018) to 30–90 days (Meyers et al., 2020). All type-2 studies monitored and reported on behavioral change of single groups. None of the type-2 studies were included in a prior meta-analysis (Gendreau et al., 2014).

Type-3 studies were characterized by (a) being exclusively conducted in the United Kingdom, (ii) with relatively small sample sizes, (iii) qualitative research methods, (iv) non-experimental RSPs and (v) non-immediate rewarding. The findings of these four recent studies (2011–2020) were all based on either semi-structured interviews, in-depth interviews or extensive ethnographic research, in prisons where the Incentives and Earned Privileges (IEP) scheme was in place. None of the type-3 studies were included in a prior meta-analysis (Gendreau et al., 2014).

3.2. Effectiveness of RSPs

Effects are discussed per study type, by focusing on target populations, target behaviors and attitudes, and techniques (including reward types), as empirical research building upon the operant-conditioning framework suggests that these variables can moderate RSP effectiveness (Andrews and Bonta, 2010; Kazdin, 1982, 2002; Premack, 1965; Schunk and Zimmerman, 2012; Sulzer-Azaroff and Mayer, 1991).

3.2.1. Type-1 studies

All but one type-1 study (McKee, 1971) researched a token economy, which is why we do not differentiate between the effectiveness of different techniques here (e.g. contracts, systems using pre-selected sets of rewards awarded upon level progression).

A review of these type-1 studies on the effects of token economies (n = 8) indicated that, overall, these RSPs were effective in improving academic achievement (e.g. language skills), and personal hygiene and cellblock maintenance, through immediately rewarding with tokens exchangeable for a range of reward types.

The synthesis of these studies also showed that the effects of these studies did not differ according to the type of target population. The effects of RSPs on individual academic achievement and performance of daily-living activities were found to be predominantly positive in volunteers who cognitively functioned below average (IQ 80–89). Positive effects on these outcome measures were found for seven out of nine studies, six of which included cognitively below average functioning (M IQ = 88.3) men aged 17–35 years who volunteered for experiment participation (Kandel et al., 1976; McKee, 1971, 1972; Milan and McKee, 1976; Milan et al., 1979a, study 1 & 2; Milan et al., 1979b), and one included adolescent girls (aged 11 to 15) with learning difficulties (Rice, 1970). Three studies reported that positive effects quickly declined to baseline level after the token economy had been terminated (Milan and McKee, 1976; Milan et al., 1979a, study 1). Notably, the sample of four studies, containing adult men, was identical (Milan and McKee, 1976; Milan et al., 1979a, study 1 & 2). There was no indication from these studies that individual characteristics (e.g. age, gender, ethnicity, sentence length) influenced the effects found. However, apart from cognitive functioning, these studies did not report data on risk,

need or responsivity of the participants.

Furthermore, the synthesis of type-1 studies also showed that the effects of the RSPs did not differ according to target behaviors and attitudes. Positive findings related to target behaviors focused on academic achievement (Kandel et al., 1976; McKee, 1971, 1972; Milan et al., 1979b; Rice, 1970), maintenance of personal and area hygiene (Milan and McKee, 1976), violent behavior (Ellis, 1993), rule compliance (Milan et al., 1979a, study 1), and time watching news (Milan et al., 1979a, study 2). Behavioral change in type-1 studies was measured either through observations (e.g. news watching), incidents reported or number of academic tests passed. Type of measurement did not appear to affect RSP effectiveness. As it was possible to exchange tokens for a wide range of reward types (e.g. material, social, internal freedom) in the RSPs reported on in these studies, and the frequency with which reward types were selected by participants was not documented, it was impossible to distinguish the effect of reward type on RSP effectiveness. Reward types were unknown in one study (McKee, 1971).

3.2.2. Type-2 studies

A review of the type-2 studies (N = 8) showed that the effectiveness of the examined RSPs differed considerably. These recent quantitative studies indicated that RSPs can be effective in advancing mental health among mentally ill participants, decreasing violent behavior among high-risk participants, increasing academic achievement, and reducing problem behavior among adolescents and young adults. The RSPs did not appear to be effective in decreasing misbehavior – both during and post imprisonment – in heterogenous populations, nor in promoting engagement in substance-use treatment among incarcerated men and women with substance use problems.

The synthesis of these studies also showed that RSP effects differed according to target population type. Although it is difficult to draw robust conclusions, effects generally were more positive for smaller and specific target populations, compared to larger and more heterogeneous groups of participants. This is illustrated by the fact that lower-quality studies on RSPs in female adolescents (aged 14 to 18) and young adults (aged 18 to 21) (in both minimum-security and maximum-security settings) (Feinstein, 2003; Sharma and Marino, 2017) and high-risk men in a restrictive housing unit (Reid et al., 2000) reported positive effects on participants' behavioral change. There were, however, three exceptions, reported in studies with somewhat higher study quality. Another study on high-risk men in restrictive housing reported mixed outcomes: no change in mental health, but a considerable decrease in antisocial behavior lasting over twelve months after RSP termination (Meyers et al., 2018). In contrast, a study on individuals with a serious mental illness found significant positive effects of RSP on mental health, but negative effects on minor violations (e.g. maintaining personal hygiene, littering, smoking) (Meyers et al., 2020). Although no clear explanation for these findings was provided, the authors did note that minor violations increased leading up to RSP participation and gradually decreased during RSP participation. Effects in this study were less favorable for both participants at higher risk of suicide or self-harm and participants refusing to partake in programmed activities more than once. Furthermore, a study among incarcerated individuals with substance use problems reported no significant change in treatment engagement and psychosocial functioning (Burdon et al., 2013). RSPs implemented in entire prison systems, and thus in heterogenous populations, illustrated less favorable results overall (Liebling, 2008; Morar et al., 2019). Moreover, effects reported in one of these studies were even more negative for compliant, at-risk-of-suicide, older and more educated individuals in the UK system (Liebling, 2008). Based on three studies including both male and female participants (Burdon et al., 2013; Liebling, 2008; Morar et al., 2019), there is no indication that effects were gender-specific.

Furthermore, synthesis also showed that the effects of the RSPs studied differed according to target behaviors and attitudes. Generally, target behaviors aimed at cognitive change, particularly dynamic criminogenic needs, were associated with increased RSP effectiveness. Specifically, seven studies included target behaviors aimed at cognitive change, of which five exclusively (Feinstein, 2003; Reid et al., 2000; Sharma and Marino, 2017) or predominantly (Meyers et al., 2018, 2020) reported positive outcomes. These cognitive programs focused on changing dynamic criminogenic needs: criminogenic cognitions, school and work, and substance use (Andrews and Bonta, 2010). Programs included participation in cognitive restructuring programs aimed at changing antisocial thinking, attitudes, behavior and choices, dealing with problems and conflict resolution (Sharma and Marino, 2017), self-reflective self-study programs, cognitive-behavioral group therapies, and need-oriented individual counseling (Meyers et al., 2018, 2020), anger management and positive peer culture (Feinstein, 2003), as well as a range of group cognitive and mental health programs addressing thinking errors, self-esteem and moral reasoning (Reid et al., 2000). However, one study included cognitive-behavioral therapy, but found contradictory results on measures of criminal thinking and psychosocial functioning, and non-significant results on substance use treatment engagement, compared to a control group (Burdon et al., 2013). It was unclear to which degree one of the RSPs included programs targeting cognitive change (Morar et al., 2019).

Finally, the synthesis of these studies illustrated that the effects of type-2 studies slightly differed according to technique. Studies show somewhat more favorable behavioral outcomes for RSPs in which pre-selected rewards were awarded upon level progression (which in some cases took up to more than a few days), compared to the use of exchangeable tokens – which, in contrast, allow for personal selection of rewards, directly after displaying a target behavior. Two studies reporting on token economies did not find significant effects on post-release recidivism and psychosocial functioning (Burdon et al., 2013; Morar et al., 2019), whereas a prior study on an RSP among incarcerated adolescents did find a decrease in problem behaviors (Feinstein, 2003). However, both outcome measures and measurement instruments differed considerably among these studies. Out of five studies using pre-selected rewards in level systems, three found merely positive results (Meyers et al., 2018; Reid et al., 2000; Sharma and Marino, 2017), one study reported mixed outcomes (Meyers et al., 2020) and another study found negative effects on behavioral outcomes (Liebling, 2008). The time between display of target behavior and reward reception did not seem to have an impact on RSP effectiveness, even though this time varied considerably between studies. In general, there were no indications that RSP effectiveness was impacted by selection of reward type. However, RSPs in which internal freedoms (transfer to general population (in restrictive housing), time out of cell, freedom to

move around the unit or facility) were used as rewards, reported more favorable outcomes regarding behavioral change compared to those that did not (Meyers et al., 2018; Reid et al., 2000; Sharma and Marino, 2017), yet not exclusively (Burdon et al., 2013).

3.2.3. Type-3 studies

Overall, type-3 studies (n = 4) indicated that RSP effectiveness can be affected by (perceptions of) program elements and implementation factors. The scope of these studies was wider than the study of RSP effects, yet contained relevant data on RSP effects. Type-3 studies were not synthesized by target population, target behaviors and technique, not least because they all concern participants in the UK's three-leveled IEP scheme. Instead, we describe the most prominent themes emerging from these studies, which can influence RSP effectiveness: the pain of self-government, perceived legitimacy, and family-contact rewards.

The synthesis of type-3 studies suggested that RSP effectiveness can be moderated by the pain of self-government. Three studies, using semi-structured interviews (Khan, 2016, 2020) and extensive ethnographic research (Crewe, 2009, 2011), reported that the pain of self-government can frustrate RSP effectiveness. This psychological burden was described as resulting from being held accountable for one's behavior (Crewe, 2011), while one's agency is simultaneously constrained (Khan, 2020). This led participants to feel constantly "on edge", as they were held "responsible for an increasing range of decisions" and experienced "less freedom to be left alone and move through the system passively" (Crewe, 2011, 518–519). This pain can frustrate RSP effectiveness either because individuals lack adequate self-regulatory skills to cope with this burden and display target behaviors, because it may trigger resistance, or both. Both incapability and demotivation to display target behaviors have been documented in studies throughout the last decade, indicating that RSP responsivity is a persistent factor (e.g. Crewe, 2011; Khan, 2020).

Furthermore, synthesis of these studies indicated that perceived legitimacy may affect RSP effectiveness. Participants in two studies (Crewe, 2011; Khan, 2016) reported that RSPs impacted their behavioral choices more positively when they felt that rewards and behavioral assessments were applied legitimately (i.e. neutral and respectful). These findings were in line with Liebling's (2008) questionnaire-based outcomes. Khan (2016) found that participants at higher system levels perceived the system as more legitimate, compared to those at lower levels. Additionally, participants who felt that they were treated respectfully by staff, chose to display target behaviors more often (Khan, 2020). In contrast, participants reporting little perceived legitimacy, also experienced decreased psychological well-being, demonstrated by feelings of powerlessness (Khan, 2016, 2020), fear and helplessness (Crewe, 2011). Contributing to a feeling of powerlessness, particularly among introvert and passively complying participants, was the experience that prison staff often failed to notice and monitor performed target behaviors (Khan, 2020). Being required to 'actively' (i.e. visibly) engage in prosocial behavior may have decreased progression opportunities for those individuals whose goodbehavior went unnoticed, which in turn demotivated them to display target behaviors (Khan, 2020). Hence, both perceptions of system application and outcomes (i.e. rewards) appear to affect RSP effectiveness, through altering extrinsic motivation levels.

Finally, synthesis of these qualitative studies highlighted that using family contact as a reward can positively influence RSP effectiveness, whilst simultaneously creating resistance. All studies (Booth, 2020; Crewe, 2011; Khan, 2016, 2020) reported that using family contact as a reward (i.e. additional visits, telephone credit; British Ministry of Justice, 2013) had an overall positive effect on incarcerated individuals' display of target behaviors, even though they generally perceived this as highly unfair. Interviews with incarcerated mothers illustrated that mother-child contact is an important social need, which serves as an important motivator for RSP compliance (Booth, 2020). The RSP was experienced as a barrier to fulfil this social need when mothers were unable to maintain contact due to a lack of telephone credit (a reward tied into level progression). Yet, contact with children was also experienced as a source of motivation for system compliance among incarcerated fathers – that is, for those attributing value to social contact (Khan, 2020). In essence, these qualitative findings not only illustrate that the attractivity of rewards is personal and subjective, but also that behavioral change can be increased by using attractive rewards. However, rationing what matters most may also create resistance and perhaps inflict upon perceptions of legitimacy.

In summary, studies on experimental token economies conducted in the 1960s–1970s found positive effects on academic achievement and hygiene maintenance, in target populations cognitively functioning below average (IQ 80–89), although statistical significance of these findings is unclear (type-1 studies). Contemporary, non-experimental RSPs yielded positive behavioral outcomes when targeting high-risk participants, including cognitive-program engagement as target behaviors, and using pre-selected sets of rewards awarded upon level progression. RSPs applied in large, heterogenous target populations were generally less effective (type-2 studies). Qualitative studies indicated that several implementation factors can influence system compliance and behavioral change among participants, such as the pain of self-government, perceived legitimacy, and using family contact as a reward. Severity of impact may depend on participant characteristics and perceptions of system application (type-3 studies).

4. Discussion

Until now, criminological literature lacked an oversight of the effectiveness of reward systems in prison (RSPs). Therefore, this systematic review set out to answer the research question: What are the effects of reward systems in prison (RSPs) on the behavior and attitudes of its participants? Twenty-one studies met the inclusion criteria, which could be categorized in three distinct types: type-1 studies concerned 20th-century research on experimental token economies, type-2 studies were contemporary, non-experimental, and quantitative, and type-3 studies were qualitative and UK-based. It stands out that while RSPs are widely implemented, the quality and quantity of recent and relevant research lags behind. Still, we can build on the synthesis of findings from each of these categories of studies.

The synthesis of type-1 studies (n = 9) illustrated that experimental token economies improved academic achievement and hygiene maintenance in target populations cognitively functioning below average. This finding is in line with the results from a meta-analysis

on correctional token economies using both punishments and rewards, in which five of these studies were included (Gendreau et al., 2014). However, the internal validity of these studies is questionable. The ABA designs (having non-contingency periods between token interventions) of most of these studies may have consolidated learning (Gendreau et al., 2014). However, the internal validity of these studies is questionable. The ABA designs (having non-contingency periods between token interventions) of most of these studies may have consolidated learning (Gendreau et al., 2014). Moreover, four of these studies were conducted on the same participants, facilitating testing effects. Furthermore, voluntary RSP participation may indicate a relatively high motivation baseline, which is generally associated with offender-treatment completion (Andrews and Bonta, 2010) and, in turn, behavioral change (Olver et al., 2011). Lastly, one third of the studies reported that positive effects deteriorated after experiment termination, indicating that token economies are suitable for managing behavior in prison but are not associated with lasting change. This corroborates findings of meta-analyses on contingency management application in non-incarcerated individuals with substance use problems (Benishek et al., 2014; Sayegh et al., 2017). A common explanation for short-lived effects of behavioral interventions is that rewarding prosocial behavior does not unlearn antisocial behavior (Crossley et al., 2013), nor instills intrinsic motivation to do so.

The synthesis of type-2 studies (n = 8) indicated that contemporary, non-experimental RSPs reported positive behavioral outcomes when targeting high-risk participants, including cognitive programs as target behaviors, and using pre-selected sets of rewards awarded contingent on level progression. RSPs applied in large, heterogenous target populations were generally less effective. RSPs including cognitive program elements adhered to multiple RNR-principles, such as addressing dynamic criminogenic needs (e.g. antisocial cognition and substance abuse) through cognitive-behavioral treatment (CBT) (e.g. Thinking for a change) (Andrews et al., 2011; Bonta and Andrews, 2007). Research on the cumulative effect of reward application and contingent CBT engagement of incarcerated individuals is scarce, and available evidence is mixed (Bahr et al., 2012; Carroll et al., 2012a, b). More importantly, effects might differ for systems paired with cognitive change programs, compared to systems in which engagement in such programs is part of an RSP and thus rewarded. The conclusion that RSPs applied in heterogenous prison populations, using universal rewards, produce heterogenous effects (e.g. Liebling, 2008), could partially be explained by the subjective valuation of reward types (Booth, 2020; Kazdin, 1982; Khan, 2016, 2020). That is, some individuals may care for (some) rewards, others may not. Lastly, somewhat more support was found for systems that use sets of rewards awarded upon level progression, compared to token economies. This is surprising, as prior research indicated that the attractiveness (Kazdin, 1982; Sulzer-Azaroff and Mayer, 1991) and immediacy of receiving rewards can moderate system success, due to reward discounting and poor impulse control, which characterize incarcerated populations (Arantes et al., 2013; Hanoch et al., 2013; Meijers et al., 2015). However, in systems that use sets of rewards, participants could not choose their rewards (which ensures rewards are sufficiently attractive) and often had to wait much longer before receiving rewards (up to 120 days), compared to token techniques.

The synthesis of type-3 studies (n = 4) suggestedthat several design and implementation factors may influence system compliance and behavioral change among participants, such as the pain of self-government, perceived legitimacy, and using family contact as a reward. While it was previously argued elsewhere that system design and deliverance can influence contingency management system outcomes (Gendreau and Listwan, 2018), these specific factors, derived from qualitative research, have not been related to RSP effectiveness before. First, being held accountable for a wide range of behaviors, but being incapable or unwilling to act responsibly, was found to frustrate system effectiveness. Qualitative studies described that an RSP can decrease psychosocial functioning of participants by over asking their self-regulatory capacities (Crewe, 2011). Individuals with an intellectual disability in prison may be especially prone to over asking, due to cognitive, social, practical, and conceptual impairments (APA, 2013). This potentially unresponsive subpopulation might be substantial, as intellectual disabilities are very common in (UK) prisons (Ali et al., 2016; Hassiotis et al., 2011; Murphy et al., 2017). We do not suggest that these individuals are unresponsive to reward systems in general, but rather to general reward systems. In fact, a long line of empirical research on motivation in students with intellectual disabilities suggests that this population is indeed highly responsive to extrinsic rewards (Katz and Cohen, 2014). However, there is a lack of empirical research on incarcerated individuals with intellectual disabilities. Second, decreased perceived legitimacy was found to decrease participants' well-being, worsen staff-prisoner relations and subvert system success of the UK's Incentives and Earned Privileges scheme (e.g. Crewe, 2011; Khan, 2020, 2016; also see Liebling, 2008). Respondents expressed feeling being punished twice when arbitrarily denied rewards; an experience which according to experimental studies can indeed deteriorate relationships between staff members and incarcerated individuals and reduce perceptions of legitimacy (Azrin and Holz, 1966; Brunton-Smith and McCarthy, 2016). In turn, poor relationships with staff could have further frustrated their progress, because maintaining good relationships with staff was also a target behavior (British Ministry of Justice, 2013). Even seminal works such as Pentonville (Morris and Morris, 1963) and Society of Captives (Sykes, 1958) already recognized that staff-prisoner relationships can have an instrumental function (both for staff and incarcerated individuals), and order and compliance are often carefully negotiated between incarcerated individuals and prison staff. Third, using family contact as a reward has been suggested to be effective for those individuals attributing high value to family contact. This is not surprising, as (incarcerated) individuals are motivated best by what they value most (Premack, 1965). Conditionalizing such basic need fulfilment also instilled resistance in this subpopulation. This is understandable, as social connectedness and autonomous decision-making are theorized to be important basic human needs (Ryan and Deci, 2000) – or 'primary goods' in criminological terms (Good Lives Model of rehabilitation, Ward and Gannon, 2006). However, the degree to which basic human-need fulfillment in prison should be behavior-based is a question on normativity rather than effectiveness, and an issue easily overlooked by criminal justice policymakers implementing behavioral interventions such as RSPs (Liebling, 2000).

5. Limitations

This study also has several limitations, mostly related to generalizability of findings of studies on RSPs. First of all, the included

studies were predominantly conducted in the UK and US, and often dated back decades. Geographical clustering makes it difficult to generalize findings to other contexts, because, next to obvious differences in prison management and characteristics between countries, institution-level factors (e.g. social prison climate, quality of staff-prisoner relationships) often differ between correctional settings, yet are also found to influence rule compliance behind bars (Bosma et al., 2020; Huebner, 2003; Pappas and Dent, 2021; Steiner and Woolredge, 2008). Geographical clustering may have been prompted by the exclusive use of English search terms - which may also have resulted in exclusion of relevant studies published in other languages. Additionally, as type-1 studies are temporally clustered in the 1960-1970s, it is questionable to what degree their findings can legitimately be used to understand effectiveness of contemporary RSPs. Prisons have changed drastically over the last 50 years. Next to specific contemporary challenges (e.g. overcrowding, gang membership, drug use, psychosocial problems), offender treatment in prison has intensified and is increasingly directed at individual risks and needs (Wooldredge and Smith, 2018). Consequentially, it is hard to generalize some of these findings over place and time.

There are also limitations to our ability to draw causal inferences related to methodological issues and study quality. Overall study quality of quantitative studies (type 1 and 2) was low in terms of SMS-scores. Poor quality ratings were mainly due to little use of control groups, checks for alternative explanations and an overall lack of randomized treatment. Effects were thus often measured as change over time in one sample. It is unclear to what extent the effects reported were caused by the RSP; perhaps behavioral change would have occurred regardless of systematic rewarding or was caused by alternative variables. As such, it is hard to pinpoint the working mechanisms of RSPs. Reported effects should therefore be interpreted with caution. However, higher and lower quality studies do not necessarily systematically differ in effect size and direction (Garrett, 1985; Pearson et al., 2002). Moreover, four type-1 studies used the same sample of volunteers, which may have overemphasized the positive effects of this study type, not least because of possible learning and testing effects, and high baseline motivation levels. It is also important to note that effects in type-1 and type-2 studies were often measured by display of target behaviors. Due to this (narrow) focus, possible side effects may have been overlooked, such as alterations in prison climate and self-regulatory capacities. Based on these findings, the rehabilitative value of RSPs is questionable, because studies often use outcome measures unrelated to criminogenic needs. This, again, raises the question to what extent display of good behavior behind bars is associated with the gradual process of psychological, moral and social rehabilitation (McNeill, 2012). On the upside, the exposed lack of high-quality empirical research can spur future empirical research on this topic and clarify the evidence base of this intervention to policymakers and practitioners.

6. Implications for theory, policy and research

This review's findings have several implications for theory, future policy and practice, and research. Mixed findings highlighted in this review might relate to a lack of adequate theory on systematic reward application in prison settings (Ward, 2019). Although operant conditioning principles are described extensively (Murphy and Lupfer, 2014), little theory is developed on how certain target behaviors and rewards are ought to motivate incarcerated individuals to change their behavior. The dominant model of offender rehabilitation (RNR) does not provide any specific answers either, other than referring back to operant conditioning principles (Andrews et al., 1990; Bourgon and Bonta, 2014) and general strain theories (Andrews et al., 2011). The RNR model appears to lack a clear conception of how extrinsic and intrinsic motivation relate to offender behavioral change (Andrews et al., 2011; Ward et al., 2006). Fortunately, the RNR model is open to "being informed by psychological models of motivation" (Andrews et al., 2011, p. 739). A next step would be to develop and test theories on RSPs, both in relation to prison-based and rehabilitative outcomes.

The poor methodological quality of many studies also point to the obvious implication of conducting rigorous evaluations of RSPs, especially considering the stakes involved. Indeed, some scholars have pointed out the problems of making programming and visitation (partly) conditional on good behavior (Craig, 2004; Hutton, 2017). Future research should therefore seek to build a strong evidence base on RSPs to inform policymaking and prison management.

At present, the evidence appears to suggest that rewards and target behaviors should be individualized and reflect participants' cognitive and other self-regulatory capacities (Bonta & Andrews, 2010; Ziv, 2017), and include most attractive rewards. The number of target behaviors should be small enough to be comprehensible by the individual (Marlowe, 2006). Target behaviors should be formulated in a measurable manner, but should also be specific and simple, because complex wording may frustrate participants (Crewe, 2011; Liebling, 2008) – especially those with impaired cognitive functioning (Gonçalves et al., 2014). Additionally, prison staff is advised to differentiate between those who are unwilling and those who are unable to take responsibility. Disappointing outcomes should not be instantly attributed to a participant failing to take responsibility. Rather, it should be examined to what degree the RSP components suit the target group, and whether the program theory needs revising (Liebling, 2008). Simple system alterations have been found to drastically alter the outcomes of contingency management interventions (see Kazdin, 1982).

This review of RSP effectiveness provides new insights, but also raises multiple new questions. First, future research could adopt search terms in multiple languages, as well as a variety of wordings, in order to retrieve as many relevant qualitative and quantitative studies on RSPs as possible. Second, the attractiveness of different reward types in different target groups could be further explored on group level, to increase participant responsivity. This could be done both qualitatively and quantitatively, as is previously done in incarcerated individuals (Gendreau and Goddard, 1991) and other target groups (Glimmerveen et al., 2018). Third, interaction effects of RSP and activities directed at cognitive change could be further explored. Increased insight in the interrelations of behavioral and cognitive change may help select effective target behaviors, suitable for decreasing prison misconduct, recidivism or other policy aims. Fourth, quantitively exploring the moderating influence of perceived legitimacy on participants' behavioral and attitudinal outcomes in RSPs would be a valuable addition to the insightful qualitative knowledge on this topic (Liebling, 2008). Finally, in congruence with the line of research on temporal discounting among incarcerated individuals, it is relevant to examine how reward gain and loss are

experienced by participants in different RSP levels. Kahneman and Tversky's (1979) seminal work suggests that reward loss will generally be valued higher than reward gain (aversion loss), which is commonly explained by negativity bias (Baumeister et al., 2001). Hence, hypothetically, effects could be more robust for incarcerated individuals who have much to lose (those in higher levels), compared to those who have little to lose (those in lower levels). As this review highlighted that good quality studies on this topic are scarce, it goes without saying that any future research should try to expose causal relations through the adoption of experimental designs.

7. Notes

1. Throughout time and disciplines, scholars and policymakers have used several terms to describe positive reinforcement, such as rewards (Howard League for Penal Reform, 2018), privileges (Chantraine, 2006), reinforcements (Milan, 1971; Glimmerveen et al., 2018), credit (Burchard, 1967; Morar et al., 2019; Durnescu and Poledna, 2020) and incentives (Liebling et al., 1999; Irish Prison Service, n.d.). In this paper, we refer to rewards to describe stimuli that are thought to increase the likelihood of future behavior occurring (Miller, 2006), as we find this lay-term more accessible than 'reinforcement', less normative than 'privilege', more general than 'credit' – which indicates a specific form of reward –, yet more specific than 'incentive', which can refer to any type of stimulus.
2. Based on the research question and inclusion criteria, we conducted a search with the following search terms: ("contingency management" OR "response cost*" OR "positiv* reinforc*" OR "negativ* reinforc*" OR incentiv* OR privileg* OR reward* OR "operant* condition*" OR "radical behav*" OR "behav* modif*") AND ("correctional institution*" OR "correctional facilit*" OR penitentiar* OR *prison* OR jail* OR incarc*). We inserted a 'NOT' command to exclude irrelevant topics: NOT ("prisoner dilemma*" OR "prisoners dilemma*" OR "prisoner's dilemma*" OR "prisoners' dilemma*").
3. No qualitative studies dating before 1993 were found. Therefore this type of study is not distinguished.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijlcj.2022.100556.

References

Ainscough, T.S., McNeill, A., Strang, J., Calder, R., Brose, L.S., 2017. Contingency management interventions for non-prescribed drug use during treatment for opiate addiction: a systematic review and meta-analysis. Drug Alcohol Depend. 178, 318–339.

Ali, A., Ghosh, S., Strydom, A., Hassiotis, A., 2016. Prisoners with intellectual disabilities and detention status. Findings from a UK cross sectional study of prisons. Res. Dev. Disabil. 53–54, 189–197.

Andrews, D.A., Bonta, J., 2010. The Psychology of Criminal Conduct, fifth ed. Matthew Bender & Company Inc, New Providence, NJ.

Andrews, D.A., Bonta, J., Wormith, J.S., 2011. The risk-need-responsivity (RNR) model: does adding the good lives model contribute to effective crime prevention? Crim. Justice Behav. 38 (7), 735–755.

Andrews, D.A., Zinger, I., Hoge, R.D., Bonta, J., Gendreau, P., Cullen, F.T., 1990. Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. Criminology 28 (3), 369–404.

American Psychiatric Association, 2013. The Diagnostic and Statistical Manual of. Mental Disorders (DSM-5). American Psychiatric Publishing Inc.

Arantes, J., Berg, M.E., Lawlor, D., Grace, R.C., 2013. Offenders have higher delay-discounting rates than non-offenders after controlling for differences in drug and alcohol abuse. Leg. Criminol. Psychol. 18 (2), 240–253.

Ajzenstadt, M., 2016. Qualitative data in systematic reviews. In: Weisburd, D., Farrington, D. (Eds.), What Works in Crime Prevention and Rehabilitation. Springer, New York, NY, pp. 237–259.

Ayllon, T., Milan, M.A., 1979. *Correctional Rehabilitation and Management: A Psychological Approach*. Wiley & Sons, Toronto.

Azrin, N.H., Holz, W.C., 1966. Punishment. In: Honig, W.K. (Ed.), Operant Behavior Areas of Research and Application. Appleton-Century-Crofts, New York, NY, pp. 213–270.

Bahr, S.J., Masters, A.L., Taylor, B.M., 2012. What works in substance abuse treatment programs for offenders? Prison J. 92 (2), 155–174.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C., Vohs, K.D., 2001. Bad is stronger than good. Rev. Gen. Psychol. 5 (4), 323–370.

Beaudry, G., Yu, R., Perry, A.E., Fazel, S., 2021. Effectiveness of psychological interventions in prison to reduce recidivism: a systematic review and meta-analysis of randomised controlled trials. Lancet Psychiatr. 8 (9), 759–773.

Benishek, L.A., Dugosh, K.L., Kirby, K.C., Matejkowski, J., Clements, N.T., Seymour, B.L., Festinger, D.S., 2014. Prize-based contingency management for the treatment of substance abusers: a meta-analysis. Addiction 109 (9), 1426–1436.

Bonta, J., Andrews, D.A., 2007. Risk-need-responsivity model for offender assessment and rehabilitation. Rehabilitation 6 (1), 1–22.

Booth, N., 2020. Disconnected: exploring provisions for mother–child telephone contact in female prisons serving England and Wales. Criminol. Crim. Justice 20 (2), 150–168.

Bosma, A.Q., C van Ginneken, E.F., Sentse, M., Palmen, H., 2020. Examining prisoner misconduct: a multilevel test using personal characteristics, prison climate, and prison environment. Crime Delinquen. 66 (4), 451–484.

Bourgon, G., Bonta, J., 2014. Reconsidering the responsibility principle: a way to move forward. Fed. Probat. 78 (2), 3–11.

Brettle, A.J., Long, A.F., 2001. Comparison of bibliographic databases for information on the rehabilitation of people with severe mental illness. Bull. Med. Libr. Assoc. 89 (4), 353–362.

British Ministry of Justice, 2011. Prison Service Instruction 11/2011. Incentives and Earned Privileges. National offender management service.

British Ministry of Justice, 2013. Prison Service Instruction 30/2013. Incentives and Earned Privileges. National offender management service.

Brunton-Smith, I., McCarthy, D.J., 2016. Prison legitimacy and procedural fairness: a multilevel examination of prisoners in England and Wales. Justice Q. JQ 33 (6), 1029–1054.

Burchard, J.D., 1967. Systematic socialization: a programmed environment for the habilitation of antisocial retardates. Psychol. Rec. 17 (4), 461–476.

Burdon, W.M., De Lore, J.S., Dang, J., Warda, U.S., Prendergast, M.L., 2013. Psychosocial functioning among inmates in prison-based drug treatment: results from Project BRITE. J. Exp. Criminol. 9 (1), 45–64.

Burdon, W.M., Prendergast, M.L., Eisen, V., Messina, N.P., 2003. Sanctions and rewards in prison-based therapeutic community treatment. Fed. Probat. 67 (2), 47–52.

Carroll, C., Booth, A., Lloyd-Jones, M., 2012a. Should we exclude inadequately reported studies from qualitative systematic reviews? An evaluation of sensitivity analyses in two case study reviews. Qual. Health Res. 22 (10), 1425–1434.

Carroll, K.M., Nich, C., LaPaglia, D.M., Peters, E.N., Easton, C.J., Petry, N.M., 2012b. Combining cognitive behavioral therapy and contingency management to enhance their effects in treating cannabis dependence: less can be more, more or less. Addiction 107 (9), 1650–1659.

Critical Appraisal Skills Programme, 2018. CASP qualitative checklist. Available at: https://casp-uk.net/casp-tools-checklists/.

Chantraine, G., 2006. La prison post-disciplinaire. [The post-disciplinary prison]. Déviance Soc. 30 (3), 273–288.

Craig, S.C., 2004. Rehabilitation versus Control: an organizational theory of prison management. Prison J. 84 (4), 928–1148.

Crewe, B., 2009. The Prisoner Society: Power, Adaptation and Social Life in an English Prison. Oxford University Press, Oxford.

Crewe, B., 2011. Depth, weight, tightness: revisiting the pains of imprisonment. Punishm. Soc. 13 (5), 509–529.

Crossley, M.J., Ashby, F.G., Maddox, W.T., 2013. Erasing the engram: the unlearning of procedural skills. J. Exp. Psychol. Gen. 142 (3), 710–741.

Cook, T.D., Campbell, D.T., 1979. Quasi-Experimentation: Design and Analysis Issues for Field Settings. Houghton Mifflin, Dallas, Ill.

Davis, D.R., Kurti, A.N., Skelly, J.M., Redner, R., White, T.J., Higgins, S.T., 2016. A review of the literature on contingency management in the treatment of substance use disorders. Prev. Med. 92, 36–46, 2009–2014.

Deci, E.L., Koestner, R., Ryan, R.M., 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol. Bull. 125 (6), 627.

Demleitner, N.V., 2017. Time off for good behavior. In: Kerley, K.R. (Ed.), The Encyclopedia of Corrections. John Wiley & Sons Inc, New York, NY, pp. 1–4.

Dixon-Woods, M., Shaw, R.L., Agarwal, S., Smith, J.A., 2004. The problem of appraising qualitative research. Qual. Saf. Health Care 13 (3), 223–225.

Durnescu, I., Poledna, S., 2020. Prison leave in Romania and the power of street level bureaucrats. Eur. J. Crim. Pol. Res. 26 (2), 231–246.

Dutch Prison Service (DJI), 2013. Handleiding Toetsingskader promoveren en degraderen. [Manual assessment framework for promotion and demotion]. Den Haag: Dienst Justitiële Inrichtingen, Ministerie van Veiligheid en Justitie.

Ellis, J., 1993. Security officer's role in reducing inmate problem behaviors: a Program Based on Contingency Management. J. Offender Rehabil. 20 (1–2), 61–72.

Emshoff, J.G., Davidson, W.S., 1987. The effect of "good time" credit on inmate behavior: a quasi-experiment. Crim. Justice Behav. 14 (3), 335–351.

Farrington, D.P., Gottfredson, D.C., Sherman, L.W., Welsh, B.C., 2003. The Maryland scientific methods scale. In: Farrington, D., MacKenzie, D.L., Sherman, L., Welsh, B.C. (Eds.), Evidence-based Crime Prevention. Routledge, London/New York, pp. 27–35.

Feinstein, S., 2003. School-wide positive behavior supports. J. Correct. Educ. 54 (4), 163–173.

French, S.A., Gendreau, P., 2006. Reducing prison misconducts: what works. Crim. Justice Behav. 33 (2), 185–218.

Garland, D., 1996. The limits of the sovereign state. Strategies of crime control in contemporary society. Br. J. Criminol. 36 (4), 445–471.

Garland, D., 2001. Mass Imprisonment: Social Causes and Consequences. Sage.

Garrett, C.J., 1985. Effects of residential treatment on adjudicated delinquents: a meta-analysis. J. Res. Crime Delinquen. 22 (4), 287–308.

Gendreau, P., Goddard, M., 1991. The realities of punishment: Implications for offender rehabilitation. University of New Brunswick, Saint John.

Gendreau, P., Listwan, S.J., 2018. Contingency management programs in corrections: Another panacea? J. Contemp. Crim. Justice 34 (1), 35–46.

Gendreau, P., Listwan, S.J., Kuhns, J.B., Exum, M.L., 2014. Making prisoners accountable: are contingency management programs the answer? Crim. Justice Behav. 41 (9), 1079–1102.

Gillin, J.L., 1931. Taming the Criminal: Adventures in Penology (No. 71). The Macmillan Company, New York, NY.

Glimmerveen, J.C., Brazil, I.A., Bulten, B.H., Maes, J.H., 2018. Uncovering naturalistic rewards and their subjective value in forensic psychiatric patients. Int. J. Forensic Ment. Health 17 (2), 154–166.

Gonçalves, P.D., Ometto, M., Sendoya, G., Lacet, C., Monteiro, L., Cunha, P.J., 2014. Neuropsychological rehabilitation of executive functions: challenges and perspectives. J. Behav. Brain Sci. 4 (1), 27–32.

Griffith, J.D., Rowan-Szal, G.A., Roark, R.R., Simpson, D.D., 2000. Contingency management in outpatient methadone treatment: a meta-analysis. Drug Alcohol Depend. 58 (1–2), 55–66.

Hall, E.A., Prendergast, M.L., Roll, J.M., Warda, U., 2009. Reinforcing abstinence and treatment participation among offenders in a drug diversion program: are vouchers effective? Crim. Justice Behav. 36 (9), 935–953.

Hamels, J.P., 1996. An Historical Analysis of Administrative and Educational Methods with Emotionally Disabled Students. Walden University, Minneapolis.

Hannah-Moffat, K., 2005. Criminogenic needs and the transformative risk subject: hybridizations of risk/need in penality. Punishm. Soc. 7 (1), 29–51.

Hanoch, Y., Rolison, J., Gummerum, M., 2013. Good things come to those who wait: time discounting differences between adult offenders and nonoffenders. Pers. Indiv. Differ. 54 (1), 128–132.

Hassiotis, A., Gazizova, D., Akinlonu, L., Bebbington, P., Meltzer, H., Strydom, A., 2011. Psychiatric morbidity in prisoners with intellectual disabilities: analysis of prison survey data for England and Wales. Br. J. Psychiatr. 199 (2), 156–157.

Higgins, J.P.T., Green, S., 2021. Cochrane Handbook for Systematic Reviews of Interventions. Available from. www.training.cochrane.org/handbook, version 6.2.

Howard League for Penal Reform, 2018. Howard League for Penal Reform's Response to the Ministry of Justice Consultation on Incentives and Earned Privileges Policy Framework. Howard League for Penal Reform.

Huebner, B.M., 2003. Administrative determinants of inmate violence: A multilevel analysis. J. Crim. Justice 31 (2), 107–117.

Hutton, M.A., 2017. Prison visits and desistance: a human rights perspective. In: New Perspectives on Desistance. Palgrave Macmillan, London, pp. 187–209.

Hyneman, C.S., 1926. Problem of Providing Incentive for Prison Labor. J. Crim. Law Criminol. 17 (4), 603–621.

Johnson, S., Stageberg, P., 2014. An Analysis on the Effects of Earned Time for Inmates Charged with Robbery. Iowa: Iowa Department of Human Rights and Division of Criminal Juvenile Justice Planning.

Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 363–391.

Kandel, H.J., Ayllon, T., Roberts, M.D., 1976. Rapid educational rehabilitation for prison inmates. Behav. Res. Ther. 14 (5), 323–331.

Katz, I., Cohen, R., 2014. Assessing autonomous motivation in students with cognitive impairment. J. Intellect. Dev. Disabil. 39 (4), 323–332.

Kazdin, A.E., 1972. Response cost: the removal of conditioned reinforcers for therapeutic change. Behav. Ther. 3 (4), 533–546.

Kazdin, A.E., 1982. The token economy: a decade later. J. Appl. Behav. Anal. 15 (3), 431–445.

Kazdin, A.E, 2002. Applied Behavior Analysis. In: Encyclopedia of Psychotherapy. 71, 94.

Keyes, D., 1996. Preventing Prison Misconduct Behavior: A Quantitative Review of the Literature. Unpublished honor's thesis, University of New Brunswick at Saint John, Canada.

Kesmodel, U.S., 2018. Cross-sectional studies–what are they good for? Acta Obstet. Gynecol. Scand. 97 (4), 388–393.

Khan, Z., 2016. An exploration of prisoners' perceptions of the Incentives and Earned Privileges (IEP) scheme: the role of legitimacy. Prison Serv. J. 227, 11–16.

Khan, Z., 2020. A typology of prisoner compliance to the Incentives and Earned Privileges scheme: theorising the neoliberal self and staff–prisoner relationships. Criminol. Crim. Justice 1–18.

Koehler, J.A., Lösel, F., Akoensi, T.D., Humphreys, D.K., 2013. A systematic review and meta-analysis on the effects of young offender treatment programs in Europe. J. Exp. Criminol. 9 (1), 19–43.

Landenberger, N.A., Lipsey, M.W., 2005. The positive effects of cognitive–behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. J. Exp. Criminol. 1 (4), 451–476.

Lee, M.S., Yu, C.T., Martin, T.L., Martin, G.L., 2010. On the relation between reinforcer efficacy and preference. J. Appl. Behav. Anal. 43 (1), 95–100.

Liebling, A., 2000. Prison officers, policing and the use of discretion. Theor. Criminol. 4 (3), 333–357.

Liebling, A., 2008. Incentives and earned privileges revisited: fairness, discretion, and the quality of prison life. J. Scand. Stud. Criminol. Crime Prev. 9 (S1), 25–41.

Liebling, A., Muir, G., Rose, G., Bottoms, A., 1999. Incentives and Earned Privileges for Prisoners: an Evaluation. Home Office, Research, Development and Statistics Directorate, Great Britain.

Lipsey, M.W., Cullen, F.T., Lipsey, M.W., Landenberger, N.A., Wilson, S.J., 2007. Annu. Rev. Law Soc. Sci. vol. 3, 297–320, 2007 (Effects of cognitive-behavioral programs for criminal offenders. The Campbell Collaboration) The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews.

Long, H.A., French, D.P., Brooks, J.M., 2020. Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. Res. Methods Med. Health Sci. 1 (1), 31–42.

Lussier, J.P., Heil, S.H., Mongeon, J.A., Badger, G.J., Higgins, S.T., 2006. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. Addiction 101 (2), 192–203.

Maconochie, A., 1859. The Mark System of Prison Discipline. Mitchell and Son Printers, London.

Marlowe, D.B., 2006. Depot naltrexone in lieu of incarceration: a behavioral analysis of coerced treatment for addicted offenders. J. Subst. Abuse Treat. 31 (2), 131–139.

Marlowe, D.B., Kirby, K.C., 1999. Effective use of sanctions in drug courts: lessons from behavioral research. Natl. Drug Court Institute Rev. 2 (1), 1–31.

McGuire, J., 2018. Implementing prison-based treatment programs. In: Woolredge, J., Smith, P. (Eds.), The Oxford Handbook of Prisons and Imprisonment, pp. 524–551.

McKee, J.M., 1971. September 11). *The Use of Contingency Management to Affect Learning Performance in Adult Institutionalized Offenders* [Meeting Presentation]. APA 1971 meeting, Washington D. C., United States.

McKee, J.M., 1972. *The Experimental Manpower Laboratory for Corrections.* Progress Report on Phase III, June through August, 1972. Rehabilitation Research Foundation.

McNeill, F., 2012. Four forms of 'offender'rehabilitation: towards an interdisciplinary perspective. Leg. Criminol. Psychol. 17 (1), 18–36.

Meijers, J., Harte, J.M., Jonker, F.A., Meynen, G., 2015. Prison brain? Executive dysfunction in prisoners. Front. Psychol. 6, 43.

Meyers, T.J., Infante, A., Wright, K.A., 2018. Addressing serious violent misconduct in prison: examining an alternative form of restrictive housing. Int. J. Offender Ther. Comp. Criminol. 6*2* (14), 4585–4608.

Meyers, T.J., Infante, A.A., Wright, K.A., 2020. Treating the seriously mentally ill in prison: an evaluation of a contingency management program. Corrections 5 (4), 256–273.

Michigan Department of Corrections, 2020. Policy Directive. Good Time Credits. Michigan Department of Corrections.

Milan, M.A., 1971. An Ecological Experiment in Corrections: A Programmed Environment for Behavior Modification.

Milan, M.A., McKee, J.M., 1976. The cellblock token economy: token reinforcement procedures in a maximum security correctional institution for adult male felons 1. J. Appl. Behav. Anal. 9 (3), 253–275.

Milan, M.A., Throckmorton, W.R., McKee, J.M., Wood, L.F., 1979a. Contingency management in a cellblock token economy: reducing rule violations and maximizing the effects of token reinforcement. Crim. Justice Behav. 6 (4), 307–325.

Milan, M.A., Wood, L.F., McKee, J.M., 1979b. Motivating academic achievement in a cellblock token economy: an elaboration of the Premack Principle. Offender Rehabil. 3 (4), 349–361.

Miller, L.K., 2006. Principles of Everyday Behavior Analysis, fourth ed. Brooks/Cole Publishing Company, Monterey, CA.

Mitchell, K., 2010. November 14). Limon prison incentive programs keep inmates in check. The Denver Post. https://www.denverpost.com/2010/11/14/limon-prison-incentive-programs-keep-inmates-in-check/.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Prisma Group., 2009. Reprint—preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Phys. Ther. 89 (9), 873–880.

Morar, I., Motiuk, L.L., Keown, L.A., 2019. Characteristics of offenders participating in the Romanian prison service credit system. Eur. J. Crime Crim. Law Crim. Justice 27 (3), 242–259.

Morgan, R.D., Flora, D.B., 2002. Group psychotherapy with incarcerated offenders: a research synthesis. Group Dynam.: Theor. Res. Practice 6 (3), 203–218.

Morris, T., Morris, P., 1963. Pentonville: A Sociological Study of an English Prison. Routledge & K. Paul, London.

Munthe-Kaas, H.M., Glenton, C., Booth, A., Noyes, J., Lewin, S., 2019. Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: first stage in the development of the CAMELOT tool. BMC Med. Res. Methodol. 19 (1), 1–13.

Murphy, G.H., Chiu, P., Triantafyllopoulou, P., Barnoux, M., Blake, E., Cooke, J., Forrester-Jones, R.V.E., Gore, N.J., Beecham, J.K., 2017. Offenders with intellectual disabilities in prison: what happens when they leave? J. Intellect. Disabil. Res. 61 (10), 957–968.

Murphy, E.S., Lupfer, G.J., 2014. Basic principles of operant conditioning. In: McSweeney, F.K., Murphy, E.S. (Eds.), The Wiley Blackwell Handbook of Operant and Classical Conditioning. Wiley Blackwell, pp. 167–194.

Olver, M.E., Stockdale, K.C., Wormith, J.S.(, 2011. A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. J. Consult. Clin. Psychol. 79 (1), 6–21.

Pappas, L.N., Dent, A.L., 2021. The 40-year debate: a meta-review on what works for juvenile offenders. J. Exp. Criminol. 152 (8), 1–30. https://doi.org/10.1007/s11292-021-09472-z.

Pearson, F.S., Lipton, D.S., Cleland, C.M., Yee, D.S., 2002. The effects of behavioral/cognitive-behavioral programs on recidivism. Crime Delinquen. 48 (3), 476–496.

Petry, N.M., 2000. A comprehensive guide to the application of contingency management procedures in clinical settings. Drug Alcohol Depend. 58 (1–2), 9–25.

Petticrew, M., Roberts, H., 2008. Systematic Reviews in the Social Sciences: A Practical Guide. John Wiley & Sons.

Premack, D., 1965. Reinforcement theory. In: Levine, D. (Ed.), Nebraska Symposium on Motivation. University of Nebraska Press, Lincoln, pp. 123–180.

Prendergast, M.L., Hall, E.A., Grossman, J., Veliz, R., Gregorio, L., Warda, U.S., Van Unen, K., Knight, C., 2015. Effectiveness of using incentives to improve parolee admission and attendance in community addiction treatment. Crim. Justice Behav. 42 (10), 1008–1031.

Prendergast, M., Podus, D., Finney, J., Greenwell, L., Roll, J., 2005. Contingency management for treatment of substance use disorders: a meta-analysis. Addiction 101 (11), 1546–1560.

Prison Reform Trust (2014). Incentives and Earned Privileges. https://prisonreformtrust.org.uk/incentives-and-earned-privileges-update-february-2014/. (Accessed 1 february 2021).

Irish Prison Service (2012). Incentivised Regimes Policy. Irish Prison Service. https://www.irishprisons.ie/about-us/care-and-rehabilitation/incentivised-regimes-policy/. (Accessed 14 May 2021).

Rawlinson, K., 2019. Prisoners could get own keys as reward for good behaviour https://www.theguardian.com/society/2019/jul/12/prisoners-may-get-perks-for-good-behaviour-in-rewards-based-reforms. (Accessed 22 february 2021).

Reid, L., Ploughe, P., Wright, R., Lehman, W., 2000. Colorado State Penitentiary Progressive Reintegration Opportunity Unit (Annotation).

Rice, R.D., 1970. Educo-therapy: a new approach to delinquent behavior. J. Learn. Disabil. 3 (1), 16–23.

Ryan, R.M., Deci, E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 55 (1), 68.

Sayegh, C.S., Huey Jr., S.J., Zara, E.J., Jhaveri, K., 2017. Follow-up treatment effects of contingency management and motivational interviewing on substance use: a meta-analysis. Psychol. Addict. Behav. 31 (4), 403–414.

Sampson, R., Glascon, S., Glen, I., Louis, C., Rosenfeldt, F., 2007. *A Roadmap to Strengthening Public Safety* (A Report to the Correctional Services of Canada Review Panel). Canada, Ottawa, Ontario.

Schunk, D.H., Zimmerman, B.J., 2012. Motivation and Self-Regulated Learning: Theory, Research, and Applications. Routledge.

Seers, K., 2015. Qualitative systematic reviews: their importance for our understanding of research relevant to pain. Br. J. Pain 9 (1), 36–40.

Serin, R.C., Hanby, L.J., 2009. Offender Incentives and Behavioural Management Strategies. Correctional Service of Canada, Ottawa.

Shammas, V.L., 2014. The pains of freedom: assessing the ambiguity of Scandinavian penal exceptionalism on Norway's Prison Island. Punishm. Soc. 16 (1), 104–123.

Sharma, D., Marino, K., 2017. Correctional management of security risk groups: a case study. J. Gang Res. 24 (4), 45–65.

Sloas, L., Murphy, A., Wooditch, A., Taxman, F.S., 2019. Assessing the use and impact of points and rewards across four federal probation districts: a contingency management approach. Vict. Offenders 14 (7), 811–831.

Steiner, B., Woolredge, J., 2008. Inmate versus environmental effects on prison rule violations. Crim. Justice Behav. 35 (4), 438–456.

Stevinson, C., Lawlor, D.A., 2004. Searching multiple databases for systematic reviews: added value or diminishing returns? Compl. Ther. Med. 12 (4), 228–232.

Skinner, B.F., 1938. The Behavior of Organisms: an Experimental Analysis. Appleton-Century, New York.

Steiner, B.M., Cain, C.M., 2014. Examining the Effects of Nebraska's Good Time Laws-Research Brief. Reports, vol. 1. Nebraska: Nebraska Center for Justice Research.

Steiner, B., Cain, C.M., 2019. The effect of removing sentencing credits on inmate misbehavior. J. Quant. Criminol. 35 (1), 89–108.

Sulzer-Azaroff, B., Mayer, G.R., 1991. Behavior Analysis for Lasting Change. Holt, Rinehart & Winston, New York, NY.

Sykes, G., 1958. Society of Captives: A Study of a Maximum Security Prison. Princeton University Press, Princeton, NJ.

Vurbic, D., Bouton, M.E., 2014. A contemporary behavioral perspective on extinction. In: McSweeney, F.K., Murphy, E.S. (Eds.), The Wiley Blackwell Handbook of Operant and Classical Conditioning. Wiley Blackwell, pp. 53–76.

Ward, T., 2019. Why theory matters in correctional psychology. Aggress. Violent Behav. 48, 36–45.

Ward, T., Gannon, T.A., 2006. Rehabilitation, etiology, and self-regulation: the comprehensive good lives model of treatment for sexual offenders. Aggress. Violent Behav. 11 (1), 77–94.

Ward, T., Polaschek, D., Beech, A.R., 2006. Theories of sexual offending. In: Brown, J., Horvath, M. (Eds.), Cambridge Handbook of Forensic Psychology, second ed. Cambridge University Press, Cambridge.

Weisburd, D., Chayet, E.S., 1989. Good time: an agenda for research. Crim. Justice Behav. 16 (2), 183–195.

Wilson, D.B., Bouffard, L.A., MacKenzie, D.L., 2005. A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. Crim. Justice Behav. 32 (2), 172–204.

Wooldredge, J., Smith, P. (Eds.), 2018. The Oxford Handbook of Prisons and Imprisonment. Oxford University Press, Oxford.

Ziv, R., 2017. The Future of Correctional Rehabilitation: Moving beyond the RNR Model and Good Lives Model Debate. Routledge.