



The Effect of Token Economy on Managing Aggression in Adults and Adolescents with Mental Disorders: A Recent Review

Fatemeh Ghojoghi ¹, Elahe Hojati Abed ^{2,*}, Mitra Khalaf Beigi ², Esmaeil Mahboubi ³

¹ School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

² Department of Occupational Therapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

³ Psychiatrist, The Founder of Omid Farda Comprehensive Treatment and Rehabilitation Center, Tehran, Iran

*Corresponding author: Department of Occupational Therapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran. Email: elahe_hojati@yahoo.com

Received 2023 November 19; Revised 2024 April 24; Accepted 2024 May 20.

Abstract

Context: Individuals with mental illness may exhibit aggression due to psychological instability. The Token Economy is a behavioral approach that aims to achieve desired changes by using tokens as reinforcements.

Objectives: This study investigates the effectiveness of token reinforcement in reducing aggressive behavior in mentally ill patients and identifies the factors that enhance the effectiveness of the Token Economy program.

Evidence Acquisition: A search for studies was conducted in the databases of PubMed, Google Scholar, ProQuest, Scopus, and Web of Science using several keywords, including "Contingency Management," "Token Reinforcement," "Token Economy," "behavioral disorder," "mental disorder," "aggression," "violence," "disruptive behavior," and "challenging behavior." The search focused on articles published from 2003 to 2023 that included the implementation of token reinforcement in various settings for managing aggression.

Results: A total of 12 studies met the inclusion criteria. After reviewing these studies, several factors were identified as responsible for the program's effectiveness. The Quality Assessment Tool for Quantitative Studies and the Oxford Centre for Evidence-Based Medicine (2011) were used to evaluate the studies. Of the 12 studies, two were rated as strong quality, one as moderate, and nine as weak quality. Additionally, three studies were classified as level 2, one as level 3, and eight as level 4, indicating significant methodological limitations.

Conclusions: Generally, 9 studies supported the effectiveness of the Token Economy in managing aggression in adults and adolescents with mental illness. This approach can be useful in the modern rehabilitation and education of individuals with aggressive behavior. We recommend further studies with higher methodological quality to examine the generalizability and transferability of the results.

Keywords: Token Economy, Mental Disorders, Aggression, Review, Behavior Therapy

1. Context

Individuals with mental health disorders, such as schizophrenia and bipolar disorder, have diagnosable psychiatric conditions that significantly interfere with their occupational performance across various domains, thereby limiting inclusion and participation (1). According to a 2021 estimate, 49.5% of adolescents aged 13 to 17 had mental disorders, and 22.8% of the adult population manifested mental disorders (2, 3). Adolescents with developmental disabilities and a genetic predisposition often face a negative social environment, leading to a high comorbidity with

mental illness. Evidence examining psychopathology and intellectual impairment found a 28% aggression rate in individuals with mental health disorders (1). This behavioral difficulty can interfere with learning new skills, maintaining attention at school, and developing effective social skills, which impedes successful interaction in various settings, such as school, and can lead to peer rejection. Consequently, adolescents may lose self-esteem, engage in alcohol and substance abuse, and become delinquent (4). Moreover, difficulty following societal rules and a tendency toward irritability and aggression can result in ongoing legal issues, potentially leading to incarceration, which

negatively affects mental and physical health and the ability to return to school or work (5). Violence and aggression towards self, others, or objects in inpatient psychiatric units can harm the individual, others, and staff, increase care expenses, and disrupt the therapeutic environment (6). When seeking community integration for adult psychiatric patients with aggressive behavior, their challenging behavior often hinders their ability to reintegrate. This results in diminished self-control, self-direction, social competencies, and a lower quality of life (7). Challenging behaviors can stem from environmental or biological factors, necessitating both pharmacological and behavioral modification methods (6, 7). These methods include de-escalation, time-outs, increased observation and support, and medication use with the patient's consent to calm them down in a non-coercive manner. If the individual does not cooperate, staff may resort to restraining the patient geographically, physically, or chemically (rapid tranquilization) as a last resort for safety reasons. However, since coercive methods can have harmful complications, particularly for mentally ill patients with physical comorbidities, and because reducing restraint and seclusion leads to shorter and fewer hospitalizations and decreased use of sedatives and hypnotics, identifying alternative behavioral plans for individuals with aggressive behavior is crucial (8). Intervention strategies known as Contingency Management (CM), based on operant conditioning, provide contingent reinforcement of exhibited desirable behavior with exchangeable tokens and have been noted to be effective in addressing behavioral difficulties (9). The Token economy (TE), a CM intervention, has historically been successfully employed as a behavior-management and motivational tool in residential, inpatient, school, and correctional settings for diverse populations. These populations include chronic psychiatric patients, students with Learning Disorders, individuals with Autism spectrum disorder (ASD), individuals with Intellectual Disability Disorder (IDD), children with Attention Deficit Hyperactivity Disorder (ADHD), individuals with substance abuse issues, and those with head injuries (10, 11). This approach has been effectively applied to various behaviors, such as reducing disruptive behavior and increasing attentive behavior at school, promoting adaptive behaviors, targeting symptomatic behaviors in inpatient settings, and improving pro-social behaviors in community-based settings (12). In 2016, a review article discussed the use of the Token economy (TE) for treating symptoms of mentally ill patients admitted to hospitals. After analyzing literature from 1999 to 2013, Glowacki et al. concluded that TE is an economically

friendly intervention that can decrease negative behaviors in inpatient psychiatric settings. They also noted that, due to the limited number of studies on the efficacy of TE for reducing negative symptoms, there is a need for further in-depth research to explore its effectiveness (13). In another review study conducted by Austin, TE was implemented to manage the behavior of students with emotional and behavioral disorders. He noted that TE is an effective intervention within the classroom; however, due to the small number of participants, there were no significant differences between pre-test and post-test in most of the reviewed studies. He recommended conducting the study again with more participants and modifications (14).

2. Objective

Therefore, this literature review aims to analyze research published within the last 20 years (2003 - 2023) that used Contingency Management in the form of TE in mentally ill adults and adolescents. The goal is to explore its impact on aggressive behavior and identify factors that increase the effectiveness of the TE program.

3. Evidence Acquisition

3.1. Search Strategy

Electronic searches were conducted by the first and second authors, limiting the results to English-language publications from the period 2003 - 2023. The search was carried out in PubMed, Google Scholar, ProQuest, Scopus, and Web of Science. The MeSH terms used included: Token Economy, Token Reinforcement, Mental Disorders, Aggression, and Violence. Other terminologies found in systematic reviews, such as Contingency Management, Behavioral Management, Psychiatric Disorders, Challenging Behavior, and Disruptive Behavior, were also applied to ensure all related articles were included. Additionally, references cited in relevant literature were examined (Figure 1).

3.2. Criteria for Including Studies

Studies including randomized controlled trials, pre-test and post-test designs, pilot studies, and case reports that employed CM in the form of TE alone or in combination with other interventions on adults and adolescents with mental, behavioral, or neurodevelopmental disorders who demonstrated aggressive behavior were included for review.

3.3. Criteria for Excluding Studies

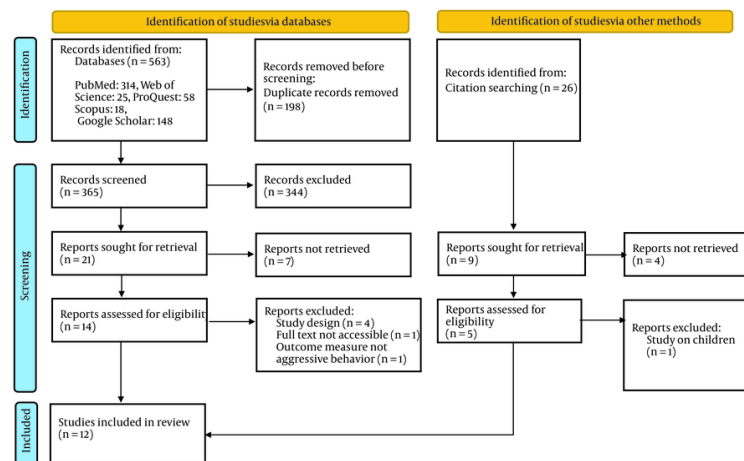


Figure 1. Identification of studies via databases/other methods

Research including child participants, specific study types such as systematic reviews, meta-analyses, anecdotal and descriptive articles, and articles without measurable outcomes were excluded.

3.4. Study Identification

Data extraction was performed by identifying relevant studies through electronic searches. Titles and abstracts were reviewed separately by the first and second authors after eliminating duplicate articles. Articles that did not meet the inclusion criteria were subsequently excluded.

3.5. Study Selection

Predetermined information such as authors, year of publication, study design, study setting, number of participants, interventions and their duration, target behaviors, results, and study limitations were screened to eliminate irrelevant studies. After combining the reviewed articles from both authors, the full manuscripts were read and assessed for final eligibility. Articles with data discrepancies were included only if both authors agreed on their importance and relevance to the purpose of this study.

3.6. Quality Assessment

The Quality Assessment Tool for Quantitative Studies (QATQS) was utilized to evaluate the quality of articles by the first and second authors. Articles were graded

based on the QATQS components of selection bias, design, blinding, methods of data collection, confounders, and withdrawals and drop-outs (15). In addition, the level of studies was determined according to the Oxford Centre for Evidence-Based Medicine Levels of Evidence (OCEBM, 2011). Articles were classified according to their design on a level of 1-5, where RCTs as the strongest evidence are considered level 1 and the weakest studies like mechanism-based reasonings are level 5 (16).

3.7. Data Synthesis

The obtained data were summarized, depicting the frequency of outcomes measured in numbers or percentages descriptively, or reporting statistical information, including average test scores, statistical significance, and effect size in changes or differences.

4. Results

Of the 589 articles identified, 570 were eliminated due to duplication and irrelevance to the criteria. This left 19 studies, of which the full texts were reviewed, except for one study that was not accessible. Data were extracted from 12 studies, each involving a reward and incentive-based Token economy (TE) for adults or adolescents over 12 years old with aggressive behavior in various settings, to determine the effectiveness of TE (Table 1). After quality assessment using QATQS, two studies were rated as strong, one as moderate, and the remaining nine as weak (Table 2). According to the

Table 2. Evaluation of the Quality of Included Articles ^a

Study	Selection Bias	Study Design	Confounders	Blinding	Data Collection Methods	Withdrawals and Drop-outs	Study Quality
Park, and Lee, (2012) (6)	2	2	1	2	1	1	1
Wolfe et al. (2003) (4)	3	2	3	2	1	1	3
Tomaka, 2009 (10)	3	2	3	2	1	1	3
Freeman and Dexter-Mazza, (2004) (24)	3	2	3	2	1	1	3
Coogan et al. (2007) (21)	3	2	3	2	1	1	3
Luby, 2007 (22)	3	2	3	2	1	1	3
Zlomke and Leland (2003) (25)	3	2	3	2	1	1	3
Holmqvist et al. (2007) (20)	2	2	1	2	1	1	1
Taylor and Mudford (2012) (17)	2	2	3	2	1	1	2
Bisconer et al. 2006 (23)	3	2	3	2	3	1	3
Meyers et al, (2018) (18)	2	2	3	2	3	1	3
Travis and Sturmey (2013) (19)	3	2	3	2	1	1	3

^a 1, strong; 2, moderate; 3, weak.

OCEBM (2011) levels of evidence, there were three level 2 studies, one level 3 study, and eight level 4 studies (Table 1). Most of the reviewed articles had low methodological quality, including very small sample sizes (N = 7) (10, 17 - 20, 23, 25), lack of long-term follow-up on the impacts of TE (N = 9) (4, 10, 17 - 20, 22 - 24), and results that were not generalizable (N = 9) (4, 6, 10, 17 - 19, 21, 23, 24). None of the studies, except for the Park and Lee study (6), included randomization and control groups. Most of the studies had only male participants, with Taylor and Mudford's study being the only one that included both sexes (17).

4.1. Token Economy (TE) in People with Mental Illness of Different Age Groups

4.1.1. Early Adolescence

Three studies involving early adolescent participants demonstrated reductions in physical and verbal aggression, property destruction, off-task behaviors, and distracting others in school settings (4, 21, 24).

4.1.2. Youth

The reduction of physical aggression and property destruction by an autistic student in the classroom did not stabilize, thus not confirming the effectiveness of TE. Additionally, disruptive talk increased after program implementation. The author suggested that the primary reason might be the student's reluctance to earn tokens in exchange for appropriate behavior (22). Token economy and self-monitoring decreased the frequency

of aggressive, disruptive, out-of-seat, and non-compliant behavior of a youth student in the school setting; however, these changes were not observed in the home group setting, which served as a control situation (25). A study by Holmqvist et al. on delinquent youths in forensic facilities compared TE plus Aggression Replacement Treatment with relational-based treatment. They noted a reduction and omission of crime (rate of sentence register & police suspension register) in both groups. The obtained results did not differ significantly between the two groups. However, the authors suggested that relationally oriented treatment might be more manageable for adolescents (20). Another study in the youth population examined the effect of a token program named Thumbs Up (signifying the positive orientation of the intervention) on behavioral statements of adolescents with mental health disorders and substance abuse. In this study, positive statements increased, but inappropriate statements (verbal expressions containing swear words, venting, and negative raving) did not change significantly because they were socially accepted by most participants during the baseline phase and remained at a low level during the token phase. The authors stated that the decreased inappropriate behavior might be due to the presence of staff in all settings (17).

4.1.3. Adults

In the adult population, three articles implemented Token economy programs in psychiatric hospitals and noted a decrease in verbal attacks, property damage,

and physical attacks toward self, staff, and other patients, which led to reduced use of interventions to control behavior (6, 23). Meyers et al. concluded that Contingency Management (CM) in chronically mentally ill prisoners was not effective in maintaining targeted behaviors like clowning and wandering (18). There were two studies on the use of TE for challenging behaviors in adults with intellectual and developmental disabilities (IDD) that indicated reductions in sexual behavior, assault, damaging objects, stripping, and pulling the siren of the emergency system, while replacement behaviors increased (10, 19). Token economy alone (6, 20, 21), TE in combination with skills training (10, 19, 23), and TE alongside self-monitoring (20, 21, 24) were the most common interventions used in the included studies. Other therapies implemented in combination with TE were level systems (10), Aggression Replacement Training (20), cooperative game therapy (4), and psychotherapy (18). Most of the studies were conducted on the adult population, with only a few studies applying Token Economy interventions to young adolescents (Figure 2).

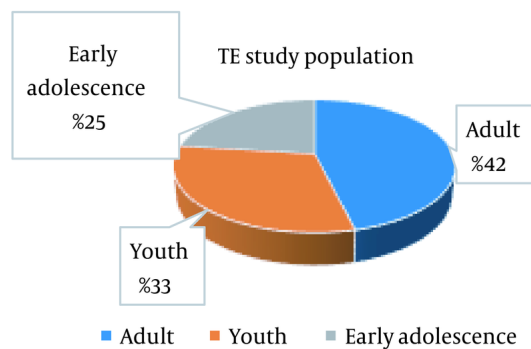


Figure 2. Chart of the Token Economy studies in different age groups

4.2. Main Factors Mentioned in Studies

4.2.1. Functional Assessment

Behavioral plans based on functional assessments can lead to meaningful reductions in aggression and significant increases in replacement behaviors. Functional assessment includes: (a) determining the target behavior; (b) identifying antecedents and situations that may stimulate the occurrence of the targeted behavior; (c) understanding the issues that arise as a result of the targeted behavior; (d) identifying

remaining functional behavior; (e) implementing environmental adaptations that decrease the targeted behavior; and (f) selecting the best strategies (23). Three studies performed individualized functional assessments for each participant with IDD and schizoaffective disorder, bipolar type (10, 19, 23).

4.2.2. Antecedent, Behavior, Consequences

The place where the challenging behavior occurs, the stimulus that triggers the target behavior, the topography of the target behavior, the timing of its occurrence, and the outcome of the behavior are all parts of the functional assessment (10). Manipulation of antecedents and environmental modifications led to significant positive changes in the three studies mentioned earlier (10, 19, 23).

4.2.3. Replacement Behavior

Differential reinforcement of alternative behaviors that encourage appropriate actions to reduce aggression has been implemented in some studies (23). Aggression Replacement Training, used in combination with Token Economy (TE) on young offenders with a history of crime, included aggression control training, interpersonal skills training, and socio-moral reasoning aimed at teaching prosocial behaviors to replace aggression. While inappropriate behavior decreased in this study, the display of prosocial behaviors was not measured (20). Taylor and Mudford observed increased positive statements, such as indicating approval, describing a desirable state or mood, empathizing, appreciating, and apologizing. However, inappropriate conversations, including swearing, venting, and negative ranting, did not significantly change from the baseline phase to the intervention phase and remained low (17). In three studies, participants avoided demonstrating undesirable behaviors while concurrently exhibiting more desirable behaviors (10, 19, 23).

4.2.4. Staff Training

Accurate staff training is an important step for the consistency of program implementation. Staff implementing the Token economy (TE) must be able to (a) identify specific settings, events, and antecedents that stimulate targeted behavior; (b) cue verbally and physically to the targeted behavior; (c) provide feedback; (d) reinforce the desired behavior contingently; and (e) offer verbal praise as a social reinforcer for desirable behavior and the absence of challenging behavior (23). All the studies included staff

training components throughout their programs. In Tomaka's study, two participants were receiving attention for opposing their peers, which reinforced them more than earning tokens for appropriate behavior. Therefore, staff began ignoring these undesirable behaviors and increased social reinforcement for exhibited replacement behaviors. This change led to a decrease in the targeted behaviors (10).

4.2.5. Inter-observer Reliability

Two independent raters simultaneously scored the observed behavioral responses. They divided the rate of their agreements by the sum of the rates of agreements and disagreements, then multiplied it by 100 to obtain the final percentage of consensus (10). Nine studies used inter-observer agreement on behavioral difficulties, with overall agreement ranging from 67% to 100% (4, 10, 17, 19-22, 24, 25). Some studies also assessed staff adherence to the TE implementation method to ensure procedural integrity and identify the need for retraining to ensure correct use of the therapy (10, 19).

4.2.6. Procedural Integrity

Because TE is a complex intervention, it is important to familiarize oneself with its principles and components and implement them precisely to achieve targeted goals (26). Monitoring staff and participants to ensure procedural integrity was conducted in four studies (10, 17, 19, 21).

4.2.7. Choosing Reinforcers

Selecting back-up reinforcers that are effective, meaningful, and age-appropriate for participants can increase the success rates and satisfaction of participants in the TE program (4). Based on the reviewed studies, rewards were determined either by the participants themselves (17, 23), by the staff's knowledge of the participants (4, 22), or cooperatively by both (10, 20).

4.2.8. Generalization

One major criticism of TE noted in the literature is the transferability and generalizability of the results to other situations (27). Taylor and Mudford extended the intervention (Thumbs Up) to new staff. When the new staff were trained, they used Thumbs Up in morning meetings of youth participants with substance abuse and mental disorders. The frequency of positive statements did not differ significantly compared to the

researcher who introduced Thumbs Up for the first time to participants, supporting its transferability to other staff (17). Travis and Sturmy assessed generalization across novel staff, antecedents, and settings. In their study, new staff members triggered five new stimulating antecedents in a non-familiar vocational and residential environment. Desirable behaviors were exhibited in the generalization phase, and none of the challenging behaviors were displayed (19). However, Zlomke and Zlomke confirmed positive behavioral changes in the classroom after TE implementation but did not observe the same results concurrently in the residential home group. They noted that they did not intend to examine the generalization of the intervention (25).

4.2.9. Social Validity

It is imperative that aspects of the study be validated by individuals beyond the primary researchers. This validation can be obtained formally or informally regarding the appropriateness of objectives, the importance of the intervention, and its effectiveness from participants, staff, or families (4). Five studies mentioned social validity, and all of them indicated satisfaction with the importance and effectiveness of TE (4, 17, 19, 21, 24).

5. Discussion

This literature review aimed to determine the effect of TE programs on the aggressive behaviors of mentally ill adults and adolescents. Evidence shows the effectiveness of TE in managing aggression in psychiatric hospitals. One study reported significant decreases in verbal and physical attacks and patient-to-staff injuries after the implementation of TE (6, 23). In the school setting, five studies were found, four of which showed reductions in aggression (disruptive, aggressive, and off-task behavior, antisocial verbal and physical behaviors) after the interventions were implemented (4, 17, 20, 21). However, in Luby's study, while property destruction and physical aggression decreased, these changes did not stabilize (22).

Token economy programs implemented in criminal settings showed moderate effects on behavior management. The number of crimes and aggressive responses decreased, and replacement responses increased, but in one study, minor violence among severely mentally ill prisoners increased with TE (18, 19, 25). In substance abuse residential care, positive statements increased, but inappropriate statements did not change from the baseline to the intervention phase

(17). Tomaka noted a decrease in inappropriate sexual behavior, physical aggression, and property destruction among home group residents with intellectual disability disorder after TE implementation (10).

5.1. Limitations and Recommendations

Based on our research, the latest study was published in 2020, which we could not access the full text of. Except for the Meyers et al. study in 2018, all other studies were conducted more than 10 years ago. Additionally, most of the studies focused on the male population, ignoring gender differences. While studies have noted that random assignment, especially with control groups, is difficult to conduct for program evaluation, considering these factors could yield different results (18). We recommend that further research with methodological improvements should be conducted on behavioral management using TE. Key recommendations include:

- (1) Implementing random assignment and including control groups.
- (2) Considering both sexes to address gender differences.
- (3) Conducting functional assessments and analyses of participants.
- (4) Providing comprehensive staff training.
- (5) Utilizing inter-rater reliability to ensure program consistency.

It is also imperative to follow up on the results to maintain the effects and ensure they can be transferred to broader societal contexts.

5.2. Conclusions

This literature review supports the efficacy of TE as a non-coercive method for reducing aggressive behaviors in adolescents and adults in various settings. The findings can be useful for establishing TE programs in modern psychiatric contexts to manage behavioral difficulties using token reinforcement effectively. This approach can help reduce injuries and the use of restrictive interventions. Although few studies have investigated the efficacy of TE for reducing challenging behaviors, most determined that TE was effective. However, TE for aggression reduction should be used with caution until further research is conducted on the topic.

Footnotes

Authors' Contribution: Conceptualization, methodology: Mitra Khalaf Beigi, and Esmaeil Mahboubi; investigation, writing-review & editing: Elahe Hojati Abed, and Fatemeh Ghojoghi; writing original draft: Fatemeh Ghojoghi; supervision: Elahe Hojati Abed. All authors gave final approval of the manuscript.

Conflict of Interests Statement: We declared that two of our authors (Elahe Hojati Abed and Mitra Khalaf Beigi, reviewer) are of the editorial boards. The journal confirmed that the authors with CoI were excluded from all review processes.

Data Availability: The dataset presented in the study is available on request from the corresponding author during submission or after publication.

Funding/Support: This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

References

1. Brown C, Stoffel VC, Munoz J. *Occupational therapy in mental health: A vision for participation*. 2nd ed. FA Davis; 2019.
2. Merikangas KR, He JP, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication-Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry*. 2010;**49**(10):980-9. [PubMed ID: 20855043]. [PubMed Central ID: PMC2946114]. <https://doi.org/10.1016/j.jaac.2010.05.017>.
3. Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: Results from the 2021 National Survey on Drug Use and Health (HHS Publication No. PEP22-07-01-005, NSDUH Series H-57). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. *J Results from the*. 2013;**2**(13):55-68.
4. Wolfe BD, Dattilo J, Gast DL. Effects of a token economy system within the context of cooperative games on social behaviors of adolescents with emotional and behavioral disorders. *Therapeutic Recreation J*. 2003;**37**(2):124-41.
5. Sarkissian C. *Do Token Economies Help Manage Behaviors of Youth?: Applying a Logical Framework to Evaluate Program Effectiveness [Masters Thesis]*. California State University, Northridge; 2017.
6. Park JS, Lee K. Modification of severe violent and aggressive behavior among psychiatric inpatients through the use of a short-term token economy. *J Korean Acad Nurs*. 2012;**42**(7):1062-9. [PubMed ID: 23377602]. <https://doi.org/10.4040/jkan.2012.42.7.1062>.
7. Tillquist K. *Reducing Challenging Behaviors through a Response Cost Token Economy in Adults with Intellectual Disability [dissertation]*. Alliant International University; 2020.
8. Davison SE. The management of violence in general psychiatry. *J Advances in Psychiatric Treatment*. 2005;**11**(5):362-70. <https://doi.org/10.1192/apt.11.5.362>.
9. Hackenberg TD. Token reinforcement: Translational research and application. *J Appl Behav Anal*. 2018;**51**(2):393-435. [PubMed ID: 29468686]. <https://doi.org/10.1002/jaba.439>.
10. Tomaka A. *Evaluation of a Level System with a Built-in TE to Decrease Inappropriate Behaviors of Individuals with Mental Retardation. [Master thesis]*. University of South Florida; 2009.

11. LePage JP, DelBen K, Pollard S, McGhee M, VanHorn L, Murphy J, et al. Reducing assaults on an acute psychiatric unit using a token economy: A 2-year follow-up. *J Behavioral Interventions: Theory Practice in Residential Community-Based Clinical Programs*. 2003;**18**(3):179-90. <https://doi.org/10.1002/bin.134>.
12. Matson JL, Estabillio JA, Matheis M. Token Economy. *Encyclopedia of Personality and Individual Differences*. 2016:1-3. https://doi.org/10.1007/978-3-319-28099-8_956-1.
13. Glowacki K, Warner G, White C. The use of a token economy for behaviour and symptom management in adult psychiatric inpatients: A critical review of the literature. *J Journal of Psychiatric Intensive Care*. 2016;**12**(2):119-27. <https://doi.org/10.20299/jpi.2016.009>.
14. Austin L. *The Effectiveness of Token Economy Interventions in Adolescents with Emotional or Behavioral Disorders [Master Thesis]*. Cloud State University; 2022.
15. Thomas BH, Ciliska D, Dobbins M, Micucci S. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs*. 2004;**1**(3):176-84. [PubMed ID: 17163895]. <https://doi.org/10.1111/j.1524-475X.2004.04006.x>.
16. Jeremy H, Iain C, Paul G, Trish G, Carl H, Alessandro L, et al. Explanation of the 2011 oxford centre for evidence-based medicine (OCEBM) levels of evidence (background document). *J Oxford Centre for Evidence-Based Medicine*. 2011.
17. Taylor SA, Mudford OC. Improving behavior in a residential service for youth in drug and alcohol rehabilitation. *J Behavioral Interventions*. 2012;**27**(3):109-28. <https://doi.org/10.1002/bin.1342>.
18. Meyers TJ, Infante AA, Wright KA. Treating the seriously mentally ill in prison: An evaluation of a contingency management program. *J Corrections*. 2020;**5**(4):256-73. <https://doi.org/10.1080/23774657.2018.1530077>.
19. Travis RW, Sturmey P. Using behavioural skills training to treat aggression in adults with mild intellectual disability in a forensic setting. *J Appl Res Intellect Disabil*. 2013;**26**(5):481-8. [PubMed ID: 23925970]. <https://doi.org/10.1111/jar.12033>.
20. Holmqvist R, Hill T, Lang A. Effects of aggression replacement training in young offender institutions. *Int J Offender Ther Comp Criminol*. 2009;**53**(1):74-92. [PubMed ID: 18162485]. <https://doi.org/10.1177/0306624X07310452>.
21. Coogan BA, Kehle TJ, Bray MA, Chafouleas SM. Group contingencies, randomization of reinforcers, and criteria for reinforcement, self-monitoring, and peer feedback on reducing inappropriate classroom behavior. *J School Psychol Quarterly*. 2007;**22**(4):540. <https://doi.org/10.1037/1045-3830.22.4.540>.
22. Luby J. *An examination of the use of the TE in reducing behaviors in an adolescent with autism spectrum disorder [Master Thesis]*. Wichita State University; 2011.
23. Bisconer SW, Green M, Mallon-Czajka J, Johnson JS. Managing aggression in a psychiatric hospital using a behaviour plan: a case study. *J Psychiatr Ment Health Nurs*. 2006;**13**(5):515-21. [PubMed ID: 16965469]. <https://doi.org/10.1111/j.1365-2850.2006.00973.x>.
24. Freeman KA, Dexter-Mazza ET. Using self-monitoring with an adolescent with disruptive classroom behavior: preliminary analysis of the role of adult feedback. *Behav Modif*. 2004;**28**(3):402-19. [PubMed ID: 15104869]. <https://doi.org/10.1177/0145445503258982>.
25. Zlomke K, Zlomke L. Token economy plus self-monitoring to reduce disruptive classroom behaviors. *Behav Anal Today*. 2003;**4**(2):177-82. <https://doi.org/10.1037/h0100117>.
26. Ivy JW, Meindl JN, Overley E, Robson KM. Token Economy: A Systematic Review of Procedural Descriptions. *Behav Modif*. 2017;**41**(5):708-37. [PubMed ID: 28423911]. <https://doi.org/10.1177/0145445517699559>.
27. Dickerson FB, Tenhula WN, Green-Paden LD. The token economy for schizophrenia: review of the literature and recommendations for future research. *Schizophr Res*. 2005;**75**(2-3):405-16. [PubMed ID: 15885531]. <https://doi.org/10.1016/j.schres.2004.08.026>.

Table 1. Data Extracted from 12 Included Studies

Author, Year	Design, Level of Evidence	Setting, Population	N	Intervention	Duration	Objectives	Result	Limitation
Meyers et al, 2018 (18)	Quasi-experimental, 2	Prison, SMI	58	CM+ psychotherapy and psychoeducational group, individual counseling	1 year based on the patients' function	Risk of self-injury, n° of requesting protective custody, decrease in crime steps, non-compliance, N of visits, minor or major violations, drug violations, N of physical attack	The N of requests for custody (M = 0.16, P < .05) and supervision (M = 0.66, P < 0.05) and the N of drug abuse decreased. The N of visits raised; But there was a slight increase in the scores of mental health condition (M = 3.63, P < 0.05) due to rises of minor violence on average between Time 1 (M = 0.54) and Time 4 (M = 1.04, P < 0.05).	Limited data available to assess program outcomes, the program is restricted to one facility, lack of control group, lack of individualized reinforcement, lack of procedural integrity
Travis, and Stormay, 2013 (19)	Single subject with ABA design, 4	Residential forensic facility, adults with schizoaffective disorder and mild IDD	3	Skills training+TE (Behavioral Skills Training)	3 weeks intervention and 3 months follow-up	Rate of aggressive and replacement responses that each participant shows after presenting 15 antecedent stimuli	During A ³ , the average percentage of aggressive responses for each participant was 54%, 71%, and 64% which decreased to 12%, 10%, and 9% during B ³ . The average percentage of replacement responses during A ³ was 34%, 15%, and 26% for each participant, and during B ³ they increased to 79%, 78%, and 82%. During the generalization phase, the mean percentage of replacement responses was 74%, 71%, and 75%. During 3 months after the intervention, the N of community trips increased from 1 community trip for all participants to 5, 5, and 2.	Possibility of increased aggressive replacement behaviors due to the presentation of the experimenter in post-treatment sessions, lack of functional analysis, no follow-up, and the participants, themselves, did not evaluate the treatment efficacy in the social validity process
Park, and Lee, 2012 (6)	Quasi-experimental, 2	hospital, SMI	44	Short-term TE	8 weeks	Verbal and physical aggression, and damaging objects (Overt Aggression Scale)	After eight weeks, the experimental group showed a significant reduction in aggressive behavior for verbal attacks, from 87% to 63%, and for physical attacks and property damage, from 91% to 78%. However, the control group's scores showed increased aggressive behaviors and had significant differences (intervention group F = 194.46, P < 0.001).	Lack of randomized reinforcement individualized reinforcement
Taylor, and Mudford, 2012 (17)	Single subject with ABAB design, 3	Residential therapeutic community, youth with substance abuse and mental health disorders	16	TE (Thumbs Up)	60 sessions	Decrease of inappropriate statements (verbal states that contain swear words, venting, neg-raving), increase of positive statements (showing approval, empathizing, describing desirable state or mood)	Positive statements increased from A ³ to B ³ for most residents (11/16). A sign test with an alpha level of 0.05 showed statistical significance (Z = 2.89, P = 0.006). All participants had a low rate of verbal statements in A ³ and did not change during thumbs up (remained at near zero levels).	Limited time of intervention, no follow-up and generalization, limited data collection due to lack of control over participants entering and leaving the program, lack of individualized reinforcer
Tomaka, 2009 (10)	Single subject with ABAB design for 2 participants and AB design for 1 participant, 4	Residential group homes, IDD with behavioral issues	3	A level system with built-in TE+ skills training	Nearly 3 months (it was different for each participant)	Frequency of problematic sexual behavior, assault, damaging objects, stripping, pulling the siren of the emergency system, and escaping	Two of the participants in an ABAB design could make it to and maintain the status of level 4 after approximately 78 and 90 days. The average replacement behavior per day for the first client was A1: 2.7, B1: 4.1, A2: 3.2, B2: 4.6, and for the second one was A1: 2.3, B1: 3.7, A2: 2.5, B2: 5.3. The third participant in an AB design who avoided performing the inappropriate behaviors, demonstrated more replacement behaviors in an average of 2.2 to 2.6 per day and made it to Level 3 after 95 days	Some vague descriptions of the topographies of the behaviors, limited time of fully training and monitoring staff implementing the level system.

Author, Year	Design, Level of Evidence	Setting, Population	N	Intervention	Duration	Objectives	Result	Limitation
Holmqvist, et al. 2007 (20)	Quasi-experimental, comparison of two interventions in two institute, 2	Adolescents in residential care	57	Aggression replacement training+ TE compared to the relationally oriented treatment model	10 weeks, twice a week and 1 hour for each session, 1 year follow-up	Rate of sentence register & the police suspension register (measured by penalty value score), admission and guilt feelings	The institutions had no significant differences in the rate of criminal activity. The values were: F = 0.69, P = 0.51 for the major sentence; F = 0.80, P = 0.46 for all sentences; F = 2.60, P = 0.09 for the major suspension; and F = 0.22, P = 0.81 for the three major suspensions. 20% of the participants had no major crime after two years; 17 out of 56 had not been sentenced, while the remaining 6 had minor crimes. The effect size of all sentences was 0.71, while the score for the three major suspensions was 0.30. After B+, 9 adolescents were not reported as a criminal, and 5 participants were reported for minor crimes.	Lack of procedural integrity, limited N of participants, no control group, and randomization
Coogan et al. 2007 (21)	Single subject with ABAB design, 4	School, adolescents with inappropriate behavior	5	Group CM + Self-Monitoring	53 days	Any behavior that involves inappropriate physical contact, speaking, hostility, playfulness, distraction, noise-making, or leaving one's seat	The average percentage of disruptive behavior in student 1 was 25% in A1, 3.07% in B1, 21.06% in A2, and decreased to 5.71% in B2. For student 2 it was 30.44% in A1, 5.96% in B1, 20.33% in A2, and 8.17% in B2. In student 3, 30.35% in A1, 6.37% in B1, 25.33% in A2, and 8.95% in B2. For student 4, 38.25% in A1, 5.81% in B1, 32.33% in A2, and 12.6% in B2. In student 5, 34% in A1, 3.96% in B1, 19.89% in A2, and decreased to 4.96% in B2. Students 1, 2, and 3 had significant effect sizes of 1.88, 1.36, and 1.30, respectively. Whereas, students 4 and 5 had the largest effect sizes (2.24 and 2.26, respectively).	Lack of follow-up
Luby, 2007 (22)	Single subject with ABA design, 4	Residential school, an adolescent with ASD	1	TE	9 weeks	Frequency of assault, damaging objects, and inappropriate statements on the behavioral datasheet	The TE was an ineffective method to modify the participant's behavior. However, assault decreased from Stage I of 0 to 10 to Stage III of 0 to 9 (the highest frequency: 49th day & the lowest frequency: 42th - 45th day), and damaging objects decreased from Stage I of 0 to 6 to Stage III of 0 to 2 (the highest frequency: 45th day & the lowest frequency: 42th - 45th day); but these behaviors did not stabilize. Disruptive talk increased from Stage I of 0 to 7 to Stage III of 0 to 15.	Single subject design, the unwillingness of the participant to earn tokens, and the antecedent stimulus that triggered the target behavior were not addressed and most of the target behaviors occurred in the classroom, lunchtime, and cafeteria
Bisconer et al. 2006 (23)	Single subject with AB ^a design, 4	Psychiatric hospital, schizoaffective bipolar type, mild IDD	1	Social skill training+ contingent reinforcement	42 months	Aggression towards others, personnel, and self, frequency of restrictive interventions usage	A reduction in the rate of aggression led to a lesser need for restrictive treatments to manage aggression and a reduction in staff injury. In addition, the TE has let the participant learn appropriate behaviors and skills, and to move to a less restrictive community home.	Lack of procedural integrity and follow-up sessions
Freeman, and Dexter-Mazza, 2004 (24)	Single subject with ABAB design, 4	Classroom of a Residential facility, ADHD, conduct disorder, adjustment disorder, and learning disorder	1	Self-monitoring, skill training+ TE	3-4 times per week, 24 sessions	Disruptive behavior and off-task behavior	Self-monitoring plus matching (adult feedback as a social reinforcer) was effective in preventing disruptive and off-task behavior in the classroom. The frequency of problem behavior decreased from 40% of intervals in the baseline to 12.06% when therapy was introduced. The mean percentage occurrence of targeted behaviors was M = 21.26 in A1, it decreased to M = 20.89 in B1, and in withdrawal and reimplementation of the matching process, they reached 28.6 and 13.2.	Significant discrepancy between the N of observations during B+, with earlier phases with a greater N of observations, no follow-up, and contingency of reward with accurate self-monitoring made it difficult to determine which variable was responsible for changes
Wolfe et al. 2003 (4)	Single subject with A-B-B1-B2-BC2 design, 4	School, adolescents with emotional and Behavioral Disorders	3	TE+ cooperative games	35 sessions based on the school year (started in the second semester and ended with the termination of the school year)	Pro-social (compliment, supporting, thanking, handshakes, or high fives) and anti-social actions (swearing, physical threats, name-calling, assault)	Prosocial verbal and physical behaviors increased during BCI and BC2. B1 and B2 had lower means compared to the means of BCI and BC2. Antisocial verbal behaviors were higher in the A and B1 phases, but in BCI these behaviors decreased. During B2 behaviors returned to baseline but with the re-initiation of the BC2 these behaviors again decreased and were virtually non-existent. During A and BCI antisocial physical behaviors were high; but when BCI was initiated	

Author, Year	Design, Level of Evidence	Setting, Population	N	Intervention	Duration	Objectives	Result	Limitation
Zlomke and Ireland 2003 (25)	Single subject with ABCB design, 4	School and residential group homes, youth with bipolar and conduct disorder	1	TE + self-monitoring	19 weeks	Decrease of minor behavior (non-compliance and being out of the seat), disruptive behavior (excessive talking out loud and singing), and aggression	these behaviors dropped to zero and remained there till the end. In the school setting, A significant reduction in challenging behaviors occurred from a baseline mean of 118 after implementing TE. The decrease continued during the TE phase (mean = 63). Introducing self-monitoring alongside the TE resulted in a significant decrease in challenging behaviors. (mean = 7.75). However, the occurrence of targeted behaviors did not change in the group home. During the maintenance phase (4 weeks) compared to the baseline, the rate of challenging behaviors remained low, in both the school and group home (mean = 13).	A small period of observation, during B2 the primary observer was not able to attend 2/3 sessions which is a concern for inter-rater agreements. Lack of procedural integrity, antecedent stimulus that triggered the target behavior were not addressed

Abbreviations: TE, token economy; CM, contingency management; SMI, severe mental illness; N, number; ADHD, attention-deficit/hyperactivity disorder; IDD, intellectual disability disorder; ASD, Autism spectrum disorder.

^a A: Baseline phase, B: Intervention phase, BC: Combination of the interventions.