Clinical Interventions in Aging

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

On the utility of within-participant research design when working with patients with neurocognitive disorders

This article was published in the following Dove Press journal:
Clinical Interventions in Aging
23 July 2015
Number of times this article has been viewed

Hanna Steinunn Steingrimsdottir
Erik Arntzen

Department of Behavioral Science, Oslo and Akershus University College, Oslo, Norway

Abstract: Within-participant research designs are frequently used within the field of behavior analysis to document changes in behavior before, during, and after treatment. The purpose of the present article is to show the utility of within-participant research designs when working with older adults with neurocognitive disorders. The reason for advocating for these types of experimental designs is that they provide valid information about whether the changes that are observed in the dependent variable are caused by manipulations of the independent variable, or whether the change may be due to other variables. We provide examples from published papers where within-participant research design has been used with patients with neurocognitive disorders. The examples vary somewhat, demonstrating possible applications. It is our suggestion that the within-participant research design may be used more often with the targeted client group than is documented in the literature at the current date.

Keywords: group design, withdrawal design, multiple-baseline design, validity, dementia, Alzheimer's disease, single-subject design

Introduction

Neurocognitive disorder (NCD) is an umbrella term for different types of diseases affecting a person's social and occupational functioning with deterioration in remembering, orienting, and attending to name some.[1] Alzheimer's disease is probably the most commonly known NCD. However, there are a number of different causes to NCD diagnoses with the different diseases that fall within the NCD category having distinct pathology. The different diseases may therefore have different effects on each individual; for example, in which cognitive domain the patient is affected (eg, memory or executive functioning). Furthermore, the disease has different effects within each person depending, among other things, on the severity of the disease.[1] Consequently, this group of people may be very heterogeneous which highlights the importance of introducing individually tailored independent variables, and the importance of being able to evaluate their effect on the target behavior on individual bases. There is in fact a general agreement that practitioners should only introduce independent variables where the effectiveness has been documented empirically.[2]

As stated by Hinojosa,[3] the term evidence-based practice refers to the fact that practitioners choose the best documented treatment for their clients. The term evidence-based practice was first used within medicine in the 1990s, but was rapidly adopted within other disciplines as well. Nowadays, there is an increased awareness of the use of empirically-supported treatments (ESTs) (also known as evidence-based treatment). Division 12 of the American Psychological Association has, for instance, published standards for

Correspondence: Hanna Steinunn Steingrimsdottir; Erik Arntzen
Department of Behavioral Science, Oslo and Akershus University College, PO Box 4 St Olavs Plass, 0130 Oslo, Norway
Email hannasteinunn@simnet.is; erik.arntzen@equivalence.net

calling a treatment an EST, emphasizing the importance of utilizing EST treatments.[4] Although it is important to be aware of the demand of using ESTs, further discussion on the topic is out of scope for this paper. However, what is interesting to note is that, as pointed out by Kazdin[5] and Hinojosa,[3] the importance of EST brings with it a certain challenge to large-group studies. For example: 1) an experimenter may be faced with a problem where there are too few participants available for randomization; 2) an experimenter may be limited to provide an answer to the experimental question he has posed, hindering the possibility of providing reliable information about which treatment can be used in the clinic setting; or 3) the independent variables that are used in large group studies may be under so tightly controlled conditions that it affects the generalizability of the results.[3,5]

We will argue in this paper that the within-subject research design is an important contribution to the development of ESTs as they allow for repeated measures of the same target behavior in one individual over some period of time, provides information about the target behavior and allows an empirical evaluation of the treatment effect on individual bases.[5] In other words, in some cases, the within-participant research design may be more suitable than group studies as it allows for continuous repeated measures of the target behavior, before, during, and after the independent variable/variables have been introduced.[6–8]

Based on the preceding discussion, the argument is made that when the goal is to change socially significant behaviors, such as reducing wandering behavior, increasing activity attendance, or improving memory in NCD patients (the dependent variables), the within-participant research design may be used to gain reliable information about the effectiveness of the independent variables on the dependent variables. Therefore, the purpose of the article is to: 1) focus on the use of within-participant research designs; and 2) provide examples from the literature where these experimental designs have been used. It is important to note that in the following discussion, the term "treatment" refers to the introduction of the independent variable/variables.[9] Also, we want to emphasize that the studies that are included below are examples used for illustrative purposes.

The research question guides the choice of the experimental design

To reach what is generally known as the "golden standard" of experimental design, the sample has to consist of a large number of participants who have been randomly selected from the population under investigation. Furthermore, the sample needs to be divided randomly between one or more experimental groups (the groups that are introduced to the independent variables) and the control group (that is not introduced to the independent variables). The results are then analyzed by using statistical methods depending on the research question posed by the experimenter. The results are considered "valid" if there is a statistical difference between the experimental groups and the control group, with *P*-values of either 0.05 or 0.01. When statistical difference is found, the experimenter is able to reject the null hypothesis and assign the difference in the dependent variable between the groups to the independent variable that was used.[10]

Needleless to say, this type of experimental design can be of great value when the goal is to identify general variables that may by effective for the population as a whole, for example, when doing meta-analyses and evaluation studies of treatment packages. For example, in a recently published review article by Woods et al[11] the authors gave an extensive overview of articles on the effect of cognitive stimulation for NCD patients. The authors found that there was a generally good effect of using cognitive stimulation with these patients. However, as the authors only reviewed randomized controlled trial studies, the information about the improvement (or lack of improvement) for each individual participant was not shown. In other words, as the results from the group studies are based on the averages of two or more experimental groups, information about the individual data are not presented.

As stated by Kazdin,

> There is no methodology that is "better" than another in some abstract sense; the methodologies are all to be viewed in the context of how they contribute to our overall goals of acquiring knowledge and our ability to use them to draw valid inferences.[5]

That being said, it is safe to say that the goal of each experimental design, whether it is a group design or within-participant design, is to rule out alternative explanations for the observed behavioral change.[12] In all cases, the goal is to eliminate threats to internal validity, such as maturation and history, so that other explanations than the introduction of the independent variable can be ruled out. As pointed out by Kazdin,[5] the results from group-design studies often omit which participant the independent variable was effective for and which participant it was not effective for. This type of information is what the within-participant research design experimenter is looking for. Therefore, as stated by Baer,[13] when applying within-participant research design, Type I

errors (falsely rejecting true null-hypothesis) are "merely worrisome." On the other hand, as there is little risk of making a Type I error, the likelihood of Type II (falsely accepting false null-hypothesis) errors increases. Again, to evaluate the weight between making either Type I or Type II errors, the experimenter needs to look at the research question in hand. If the research question is to find the functional relation between the independent and dependent variable, which would be of utmost importance for the individual that receives the treatment, making Type I error should be avoided. By avoiding such error, the experimenter learns more about few important variables, variables that are "typically more powerful, general, dependable, and – very important – sometimes actionable".[13]

For that reason, in some cases, knowing more about each individual and how each individual responds to the different variables that are manipulated in an experimental setting can be more valuable than knowing about the general findings across a larger group of persons.[12–14] In order to enhance the quality of life for each individual, the independent variables need to be evaluated for that person and not based on averages from groups. Gathering information about independent variables and studying the effect on the dependent variable/variables in each individual allows: 1) continued use of the independent variables that are effective; 2) that when there is no clinically significant change in the dependent variable, the independent variables may be changed to maximize the efficiency; or 3) that if the independent variable/variables do not have any documented effect on the dependent variable, they can be removed and another independent variable introduced. Thus, the choice of the experimental design should always depend on the research question or the goal of the treatment.

Case studies, within-participant design, and within-participant research design

Before introducing the different types of within-participant research designs that can be used, it is important to take a look at both case studies and within-participant design to show how they are different from within-participant research design.

Case studies

Kazdin[5] has listed some of the key features of case studies as follows: 1) a study where one case (can by anything from one person to larger organizations or countries) is studied intensively; 2) the information that is retrieved does not focus on dependent measures, they are rather narrative reports; 3) the complexity and the nuance of the case are highlighted; and 4) although the study starts with the current situations, retrospective information are often included.

Case studies are often based on anecdotal information and they do not control for the effect of other variables that might have effect in the study, and they are seldom considered to have sufficient generalizability. In other words, there is rarely a manipulation of independent variables where the effect on the dependent variable/variables is studied,[15] there is no systematic assessment, and the study can be highly biased by the subjective experience of the experimenter. Taking these features into account, this type of study lacks the experimental control that experimenters strive for and does not allow for a valid inference to be drawn. However, it is important to consider the benefits these types of studies can have as they often report on uncommon cases where the detailed level of information can be an inspiration to new research questions and/or hypotheses.

An example of a case study is a study where deterioration in language function in a patient diagnosed with NCD was studied. The goal was to further the understanding of the breakdown in semantic knowledge.[16] The authors used different tasks to map the language deficits and could document, for example, the "progressive breakdown in referential specificity".[16] In the general conclusion of the article, the authors state that NCD patients are not a homogeneous group and there is great differences in terms of which language deficit is detected, even between patients with the same clinical diagnoses. These results are important as they stress the importance of choosing carefully, on an individual basis, the independent variables that are to be introduced for each target behavior in each NCD patient.

Within-participant design

The second type of design mentioned is within-participant design, which is also different from within-participant research design in a number of ways. This type of experimental design has few participants, and each participant may be exposed to number of independent variables. However, the repeated measures of the individual participant are discarded, and the results are presented as a type of small group design.

An example of within-participant design study is a study where the goal was to retrain activities of daily living (such as making a tea or coffee, using a CD player, or changing batteries in a remote control) in 14 NCD patients (mini–mental state examination ranging from 10 to 26).[17] There were three types of training conditions in the study: errorless learning,

modeling, and trial and error. Each condition lasted for 1 week, and the three conditions were counterbalanced for each participant during the 3-week period. The results from the study were taken together to form averages across participants, and showed that the errorless learning condition and the modeling condition had the greatest effect for the participants. The authors called for a replication of the findings using a between-group randomized controlled trial experimental design to replicate the findings. However, doing so will mask possible individual differences in learning as different procedures may suit some persons and not others. Hence, we would like to point out that this experiment may also be done by using within-participant research design, allowing the experimenters to determine the effect of the independent variable (condition) on the dependent variable (the behavior of the participant). By doing so, the experimenters would gain important information about which independent variables would be most effective for each participant by showing the effect of the independent variable on the dependent variable and excluding other possible explanations that the change in the dependent variable may be assigned to.

Within-participant research design

Within-participant research designs are very different from case studies as the goal is to minimize threats to internal validity and identify functional relation between the independent variable and the dependent variable.[18] There are different types of within-participant designs experimenters can choose from. Importantly, they all have the common characteristic of the repeated measures of the target behavior within each individual participant. The most commonly known within-participant research designs are the withdrawal design, multiple-baseline design, multiple-treatment design, and changing-criterion design. These designs have been used in some studies with patients with NCD (Table 1).

Withdrawal design; ABAB

Before discussing the withdrawal design it is important to note that withdrawal design is different from reversal design. In a withdrawal design, the independent variable is introduced and removed in the different experimental phases. In the reversal design, the independent variable is presented to one dependent variable at one time, and later to a different, incompatible dependent variable.[19] In spite of these differences, it is common to see the term reversal design when the term withdrawal design should actually be used.[20]

In a withdrawal design, there has to be a minimum of three experimental conditions (ABA).[5] However, it is preferable to end the treatment with the independent variable in effect, and therefore, the following example shows an ABAB design (Figure 1). In this hypothetical example, the rate of responding during both baseline condition (A) and during introduction of independent variable (B) is shown.

While gathering data during the first baseline condition (first A condition in Figure 1), the experimenter obtains two

Table 1 Examples of studies using within-participant research design with NCD patients. The examples are in chronological order

Author	Year	Participants	MMSE	Dependent variable	Independent variable	Design
Hussian[36]	1988	5 NCD patients	Moderate to severe	Inappropriate toileting, bed misidentification, exit seeking, inappropriate entry, ward disorientation	Verbal, physical, attentional, and perceptual cues	Withdrawal
Bourgeois[22]	1993	6 NCD patients	8–19	Conversation skills	Memory aid	Withdrawal
Heard and Watson[24]	1999	4 NCD patients	N/A	Wandering	DRO	Withdrawal
Nolan et al[37]	2001	3 NCD patients	Average 5.7	Room finding	Memory aids	Multiple baseline across participants
Altus et al[26]	2002	5 NCD patients	3–16	Communication and mealtime participation	Family style meals or prepared plates	Withdrawal
Baker et al[25]	2006	1 NCD patient	N/A	Aggression	Noncontingent escape	Withdrawal
Sellers[39]	2006	4 NCD patients	3–16	Social and agitated behaviors	Animal assisted therapy	Withdrawal
Dwyer-Moore and Dixon[38]	2007	3 NCD patients	N/A	Disruptive vocalization, wandering/exiting	DRO, NCA, and FCT + extinction	Multi-element design
Arntzen et al[23]	2013	1 NCD patient	20	Identity matching	Length of delay	Withdrawal

Abbreviations: MMSE, mini–mental state examination; N/A, not applicable; NCD, neurocognitive disorder; DRO, differential reinforcement of other behavior; NCA, noncontingent access to attention; FCT, functional communication training.

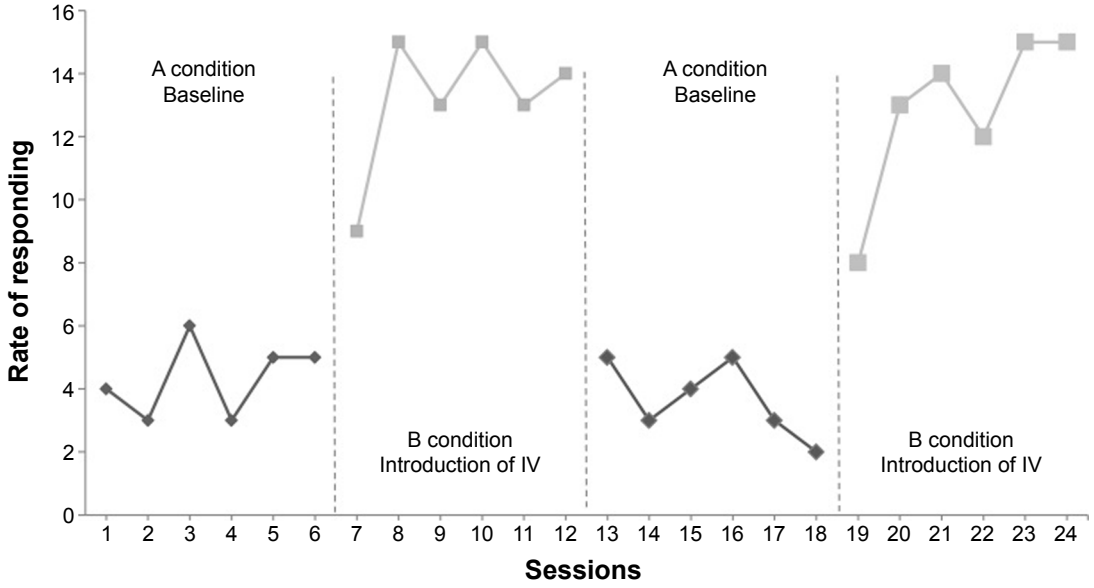

Figure 1 Hypothetical example of the withdrawal design (ABAB).
Notes: Y-axis shows number of responses per 10 minutes. The first condition is a baseline condition with low rate of responding. The second condition is the first introduction of IV where a change in the dependent variable is observed. The third condition is a reintroduction of the baseline condition where rate of responding decreases again. And the last condition is a reintroduction of the IV.
Abbreviations: IV, independent variable; A, baseline; B, introduction of IV.

types of information: descriptive information and predictive.[5] The descriptive information is about the variability and the pattern of responding. The data that are gathered during the baseline condition need to be stable (absence of a trend), or with a trend that is in the opposite direction to the expected effect of the independent variables. If the baseline data show a trend that is in the same direction as the presumed effect of the independent variables, it poses a threat to internal validity. Furthermore, the baseline needs to be stable, with as little variability as possible, for the sake of validity. The predictive information that is obtained after having documented the frequency of behavior during the baseline period provides an idea of the likelihood of the behavior being emitted again in the future, if the independent variables would not be introduced.

Following the baseline condition is a condition where the independent variables are introduced (the B condition). As for the baseline condition, this condition provides two types of information: descriptive and predictive. The repeated measures over time show whether there is a change in the target behavior in the desired direction or not. Additionally, this condition provides information about whether there is an observed difference in responding in comparison to what was predicted by the baseline condition. This is also known as counterfactual effect or counterfactual reasoning.[10]

As already noted, there needs to be at least three experimental conditions (ABA), and therefore it is important that the experimenter does not stop at this point in his study (after AB). The reason is that the observed change in behavior may have been due to other variables that were not controlled for, such as maturation or history.[5] Therefore, to eliminate threats to internal validity, the experimenter returns to baseline condition again (second A condition) by withdrawing the independent variable/variables while continuing to record the dependent variable over a period of time. Again, the second condition provides two types of information: 1) it describes changes in the target behavior; and 2) it predicts how the behavior will be if the independent variable/variables condition is not reintroduced. If the behavior goes back to the initial baseline level as it was during the first baseline condition, the likelihood of the experimenter having identified the cause–effect relationship is strengthened. At this time, the experimenter would reintroduce the independent variables (Figure 1, second B condition) and continue recording changes in the target behavior to see whether responding returns to the same level as in the previous B condition, where the independent variable/variables were introduced. If so, the experimenter can state with a higher degree of certainty that the change in the dependent variable was due to the introduction and removal of the independent variable.

The withdrawal design has been used in a number of studies to identify variables that affect the behavior of NCD patients. For example, the withdrawal design was used to increase attendance and engagement of six NCD patients in different

activities by using descriptive prompts (Table 1).[21] During the baseline (first A condition), the experimenters recorded the dependent variable: presence and engagement. The first baseline recording showed that presence was generally low, and engagement decreased through the course of the session. When introducing the descriptive prompts (the independent variable) in the first B condition of the study, there was a dramatic increase in both presence and engagement, going from 17% presence in the baseline condition to 75% in the first treatment condition, and from 78% engagement during baseline to 92% engagement in the treatment condition. The authors then removed the descriptive prompt and thereby reverted to the baseline condition again (second A condition). Returning to baseline resulted in decrease in both presence and engagement, down to 17% and 44%, respectively. In the last B condition of the study, the descriptive prompts were reintroduced, resulting in an increase of both presence (69%) and engagement (79%). Hence, the experimenters documented high experimental control where they showed consistent changes in the dependent variable (presence and engagement) depending upon the independent variable (descriptive prompts). What is worth noting here is that the data that are provided in the study are aggregated data; however, the data for each individual can be obtained from the authors.

Other studies using withdrawal design have targeted conversation skills,[22] identity matching-to-sample (MTS),[23] wandering,[24] aggression,[25] social and agitated behavior,[39] and mealtime participation[26] (Table 1). However, it is important to note that if the behavior change needs to be done quickly (such as when working with self-injurious or aggressive behavior), other research designs may be more suitable.

Multiple-baseline design

The multiple-baseline design is a type of experimental design where the independent variable/variables are implemented across behaviors, situations, or individuals. As for the withdrawal design, the multiple baseline consists of A and B conditions. Baseline performance is recorded during the A condition, providing both descriptive and predictive information about the behavior, whereas the B condition provides information about changes in the behavior dependent on implementation of the independent variable/variables (Figure 2). The introduction of the independent variable in the B conditions is the same regardless of how the design is arranged (behaviors, situations, or individuals). A major difference from the withdrawal design is that the experimenter does not withdraw the independent variable/variables as when using the ABAB experimental design. Therefore, the multiple-baseline design is in some cases better suited than the withdrawal design, as there are situations in which it may be either unethical or even impossible to withdraw the effect of the independent variable/variables.

When using multiple baselines, for example, across different kinds of behavior, the experimenter gathers data on three or more baselines (eg, Behavior 1, Behavior 2, and Behavior 3 in Figure 2). When the baseline has been recorded for sufficient stability in the data to be shown, the independent variable is introduced contingent upon Behavior 1. Meanwhile, the experimenter continues to record Behavior 2 and Behavior 3. Experimental control is demonstrated when changes are documented in Behavior 1 after introduction of the independent variable, while continuing to show stable data on Behavior 2 and Behavior 3. When stability in the data is reached for Behavior 1, the independent variable is applied to Behavior 2 (Figure 2). If the change in Behavior 2 occurs after implementation of the independent variable, and the change is in accordance with the change in Behavior 1, the experimenter can conclude that the changes in the behavior are due to the independent variable that was introduced and exclude the possibility of the effect of other variables. If the additional replication with Behavior 3 shows the same pattern of responding, the experimenter can conclude with great confidence that the change in the dependent variable was due to the introduction of the independent variable.

The multiple-baseline design across settings (work shifts) was used to study activity engagement in NCD patients living at an assisted-living facility.[27] The independent variable was a check in procedure. More specifically, a certified nursing assistant was trained to: 1) check in with each resident within 15-minute intervals; 2) provide praise on specific behaviors when the residents were engaging appropriately in an activity; and 3) make an offer to the resident of two or more activities if the resident was not engaging in any. The dependent variable was activity engagement.[27] There were five participants in the study. The authors presented the results for: 1) two residents separately; and 2) aggregated data for all residents. For the first participant, the data showed that there was a decrease in activity engagement across sessions before the implementation of the independent variables on the morning shift (during baseline). The baseline results during the night shift showed a similar pattern as the baseline for the morning shift. Following introduction of the independent variable during the morning shift, there was an increase in the dependent variable, whereas the baseline recording continued without introduction of independent variable for

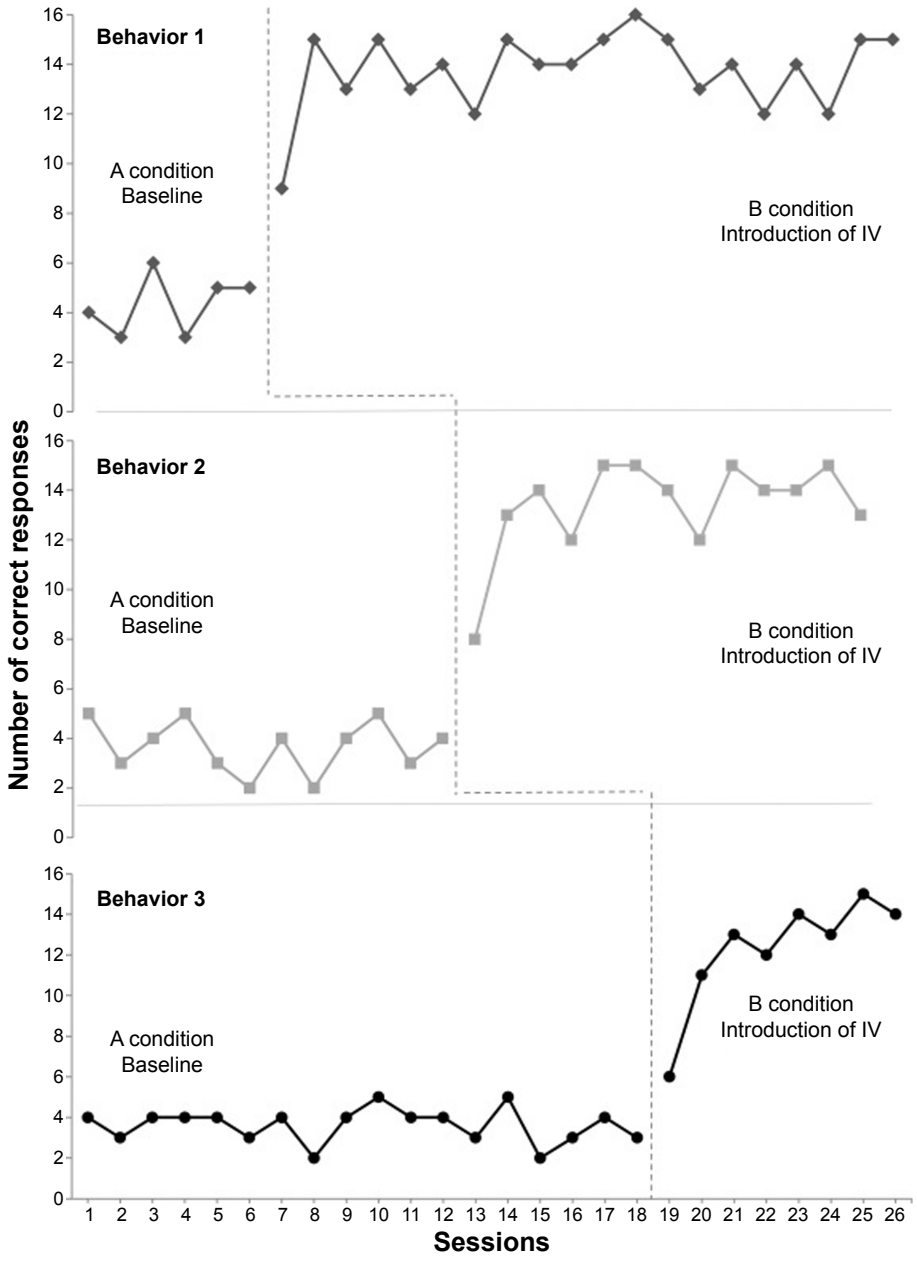

Figure 2 Hypothetical example of multiple-baseline registration across three behaviors.
Notes: Top: baseline is registered in six sessions before introduction of IV. Middle: the baseline registration continues for an additional six sessions for the second behavior before the IV is presented. Bottom: baseline registration continued for an additional six sessions before the IV was applied to the third behavior.
Abbreviation: IV, independent variable.

the evening shift. The baseline for the evening shift showed that, despite introducing the independent variable during the morning shift, it did not affect the behavior during the evening shift. When the independent variable was introduced at the evening shift as well, the behavior changed as a function of its introduction. Similar results were obtained from the second participant. The cumulative data for all participants also showed the same pattern: little activity engagement in the absence of the independent variable with an increase in the dependent variable following introduction of the independent variable, and not otherwise.

Although the multiple-baseline design may be suitable when, for example, it is either unethical or unpractical to use the withdrawal design, it is not without limitations. For example, using the multiple-baseline design can be quite time consuming and, notably, it may be inappropriate to keep one or more behaviors on hold for a long period of time. For example, as seen in Figure 2, Behavior 3 is registered

frequently for a long period of time while studying the effect of the independent variable on Behaviors 1 and 2. Therefore, other variations of the design may be preferred, such as using a nonconcurrent multiple-baseline design[28] or multiple probe design.[29] For further variations of multiple-baseline designs see, for example, Kazdin.[5]

Multiple-treatment design

The multiple-treatment design is different from the multiple-baseline design in that the former allows for two or more independent variables to be evaluated simultaneously whereas the latter targets only one independent variable at a time.[5] By introducing the independent variable during the same condition, the experimenter can exclude the effect of order such as would appear when using ABAB design. It also allows the experimenter to save time as it may be less time consuming compared to when the independent variable is introduced consecutively.

As for the other experimental designs that have already been discussed, the experimenter starts out by taking baseline data of the target behavior (Figure 3). When stability in the data has been reached, two or more independent variables are implemented. Both (or all) independent variables are presented in the B condition. However, there are some important issues that need to be addressed. First, although the independent variables are implemented in the same condition, they cannot be simultaneously in effect. As Kazdin[5] puts it, the independent variables "must 'take turns' in terms of when they are applied".[5] Furthermore, the target behavior needs to be of a kind that is possible to change rapidly between the independent variables, and additionally, it has to be possible to discriminate between the independent variables.

Multi-element design and alternating treatment design are two variations of the multiple-treatment design. In the multi-element design, the independent variables are correlated with some stimulus context, such as that one teacher introduces one type of independent variable whereas another teacher introduces another.[30] Consistent changes in the dependent variable in accordance with the independent variables that are presented would show the effect of the independent variable on the dependent variable. In the alternating treatment design, an additional experimental condition is added where the independent variable that proved to be the most effective is presented alone.

In a study by Runci et al[31] the authors discuss the importance of taking both patients' first and second language into an account when designing a treatment for NCD patients. This is because the disease may affect the two (or more) languages differently, with a possible earlier onset of deterioration in the patient's second language. Therefore, the alternating treatment design was used to help determine whether it was more efficient to use their participant's first or second language to reduce inappropriate vocalization. The participant in their study was diagnosed with severe NCD. Her first language was Italian, but she lived at an English-speaking nursing home. The authors started out by

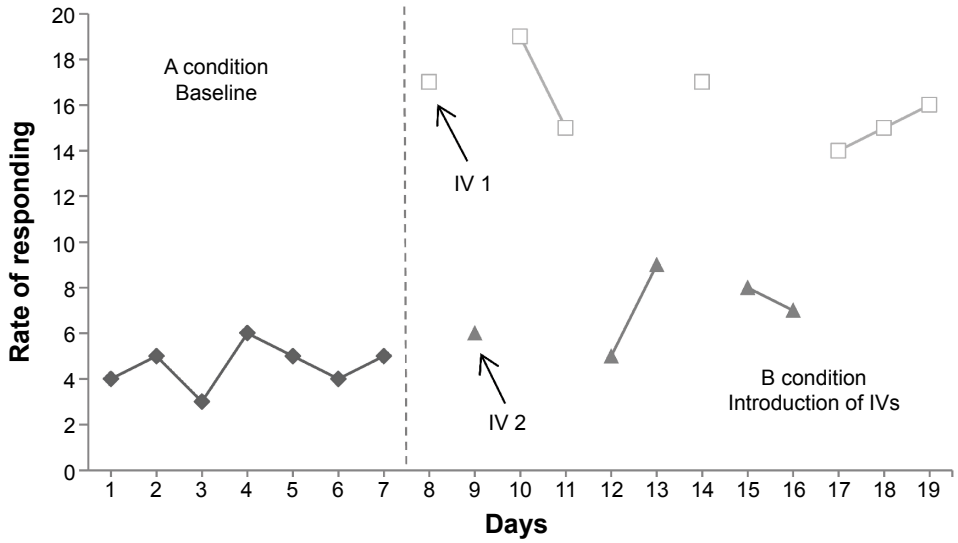

Figure 3 A hypothetical example of a multiple-treatment design is shown.
Notes: Y-axis shows number of responses per 10 minutes. Left: baseline registration of the target behavior. Right: two different IVs are introduced (IV 1 and IV 2) and the rate of responses for the target behavior is measured in the presence of one or the other.
Abbreviation: IV, independent variable.

taking baseline data of the dependent variable (noisemaking) for 10 days, along with additional registration of the antecedents and consequences of the behavior to learn about its function, that is, under what circumstances the behavior happened and consequences that followed.[31] There were two independent variables: 1) music in English and interaction with the researcher speaking English; and 2) music in Italian and interaction with the researcher speaking Italian. The results from the study showed that the participants showed more verbally disruptive behavior when in the presence of the first independent variable (English) compared to the second (Italian). Furthermore, there was a general improvement in appropriate speaking with greater effect when exposed to Italian. Thus, the authors concluded that there was a significant difference between the two independent variables with more verbal disruption in the presence of the first independent variable (English) compared to the second (Italian).

Changing-criterion design

In the changing-criterion design, following a baseline phase where the target behavior is recorded, the independent variable/variables are implemented for a gradual change in the target behavior (Figure 4).[5,32] This experimental design is different from the ABAB design as there is no need to withdraw the independent variables at any time, and is different from the multiple baselines as there is only one target behavior and no need to withhold the independent variable/variables.

The requirements for this type of experimental design are the same as for the other experimental design where there has to be stability in baseline measures before introducing the independent variable. Experimental control is assessed depending on the correlation between changes in the independent variable followed by changes in the dependent variable (Figure 4). Notably, there is no restriction of the independent variable related to the dimension of the target behavior. For example, if the dependent variable is reducing cigarette smoking, the participant should have free access to cigarettes and not only to the number of cigarettes specified by the different steps in the procedure. The changing-criterion design can be used either to increase or decrease behavior. There are to our knowledge no published studies where changing-criterion design is used with NCD patients for introduction of independent variables. Therefore, although the following example is not with participants with NCD, it is included to demonstrate how this type of experimental design could be applied.

In this example, the changing-criterion design was used for assessment of the functional relationship between independent variables (a behavioral package) and a dependent variable (aerobic exercise behavior) in five adult participants with vascular headaches.[33] Aerobic training was measured in Cooper points.[33] The results showed a functional relation between the independent variable and the dependent variable where there was a gradual increase in Cooper points for all five participants, from 0.8 Cooper points on average per week during the baseline to an average of 23.6 Cooper points per week following the introduction of the independent variable. So far the most commonly used within-subject research designs have been discussed with reference to sample of published articles from the literature. Before summing up

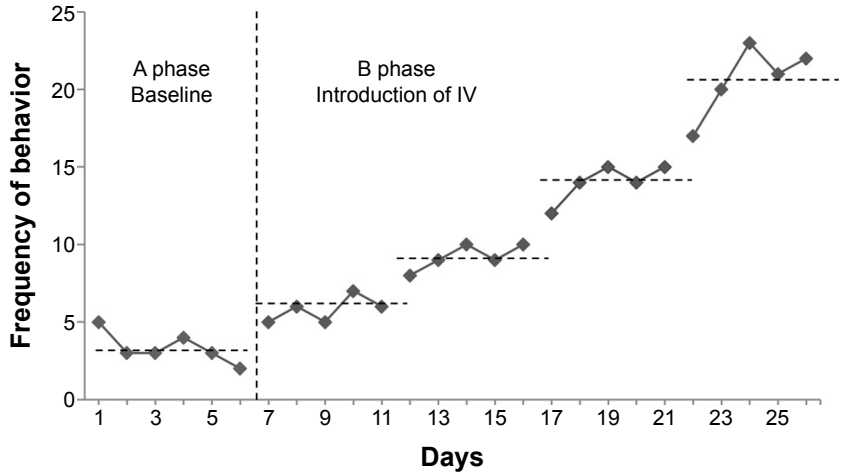

Figure 4 Hypothetical example of the step-wise change in frequency in behavior when using the changing-criterion design.
Notes: The baseline is recorded for 6 days. On day 7, the IV is introduced.
Abbreviation: IV, independent variable.

in conclusion, it is important to note two important issues, reliability and validity.

Reliability

The term reliability can either refer to the reliability of the test (common in test psychology) or the reliability of the measure employed. When using within-participant research design, the independent variables that are introduced, and their effect on the dependent variable, is studied for each individual. Consequently, reliability in within-participant research design refers to consistency in scoring of a target behavior between two independent observers.[5]

Identification of the consistency in scoring between two independent observers is gained by doing interrater or interobserver agreement.[5] In order to obtain high interobserver agreement, the target behavior must be very well defined. With proper operationalization of the target behavior, the experimenters and clinicians can agree on when the target behavior is emitted and when it is not. There are different ways of calculating interobserver agreement. One way is to do a simultaneous recording of the target behavior by two independent observers during some specified period of time. When the recording is done, the observers can calculate the number of agreements and divide it by a number of agreements plus disagreements and multiply by 100 (point-by-point agreement). That leaves the observers with a percentage score that shows how consistent they are in measurement of the target behavior. There are a number of variables that may affect the agreement, such as a number of data points. Readers are directed to Kazdin[5] for further elaboration on that. However, in general, the higher the agreement, the more reliable the observation, and therefore, agreement should be 90% or higher.

Validity

As already mentioned, the within-participant research design provides high internal validity. However, it has been stated that when increasing the internal validity of a study, it comes at the cost of the external validity.[10] Hence, although the within-participant research designs provide high internal validity, they have been criticized for low external validity. Nonetheless, it is worth noting that, as stated by Kazdin,[5] high internal validity does not necessarily mean that the external validity needs to be low. He points out that, for example, studies targeting child therapy have shown fine external validity despite the more controlled conditions in the experiments. Furthermore, as pointed out by Sidman,[34] replication of a study or a treatment plays an important role in the establishment of generalization. As stated by him, "failure to replicate a finding … is a result of incomplete understanding of the controlling variables".[34] This may sound a bit harsh, but it bears with it a positive view that once the behavior has been registered during baseline, and the counterfactual effect has been evaluated during introduction of the independent variable condition, it is possible to do an evaluation of whether correct variables were identified. Assuming the study has high internal validity, generality of the findings can then be studied by doing replications of the study. If there is a high consistency between replications, external validity is increased.[35] In sum, with successful replication of within-participant research design experiments, sufficient degrees of both internal and external validity can be accomplished.

Conclusion

The purpose of a treatment should always guide which experimental designs are used, and not the opposite. It is often the case, both in the applied setting and when doing experiments with older adults with NCD, that the treatment is intended to change certain behavior for that individual. Sometimes the goal is to increase the frequency of certain behavior and at other times to decrease it. In such cases, the use of within-participant research design is of great importance as the effect of the independent variables can be evaluated effectively. Needless to say, this is of great ethical importance. Therefore, the use of within-participant research design should be a preferred method of choice when the goal is to modify socially significant behavior as these types of experimental designs provide reliable results on cause–effect (or functional) relations. They allow for individual differences in terms of defining target behavior and to which degree the behavior needs to/should change. In this article, examples from the literature were provided for demonstration of the use of different types of within-participant research design. However, it is worth noting that the list of different within-participant experimental research designs has not been comprehensively covered in this paper.

In sum, within-participant design is a well-established research method allowing the researchers to determine the effect of the independent variable/variables on the dependent variable/variables by excluding other possible causes for the change in the dependent variable. When deciding which independent variables should be introduced when working with NCD patients, it is of utmost importance to take individual differences into an account. The conclusion is made that within-participant research design provides reliable and valid information about the effect of the independent variable/

variables on the dependent variable for each individual participant, and is thereby an effective tool for evaluating whether to continue, withdraw, or make other changes to the independent variable, all on an individual basis. It is our hope that this article has cast light on the use of within-participant design and perhaps sparks interest for experimenters and practitioners to learn more about the use of these types of experimental designs when working with NCD patients.

Disclosure

The authors report no conflicts of interest in this work.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 5th ed. Washington, DC: American Psychiatric Association; 2013.
2. Newsom C, Hovanitz CA. The nature and value of empirically validated interventions. In: Jacobson JW, Foxx RM, Mulick JA, editors. *Controversial Therapies for Developmental Disabilities: Fad, Fashion, and Science in Professional Practice*. CRC Press; 2005.
3. Hinojosa J. The evidence-based paradox. *Am J Occup Ther*. 2013;67:e18–e23.
4. Chambless DL, Ollendick TH. Empirically supported psychological interventions: controversies and evidence. *Annu Rev Psychol*. 2001;52:685–716.
5. Kazdin AE. *Single-Case Research Designs: Methods for Clinical and Applied Settings*. 2nd ed. New York: Oxford University Press; 2011.
6. Iversen IH. Single-Case Research Methods: An Overview. In: Madden GJ, Dube WV, Hackenberg TD, Hanley GP, Lattal KA, editors. *APA handbook of behavior analysis: Vol 1. Methods and principles*. Washington, DC: American Psychological Association; 2013:3–32.
7. Perone M, Hursh DE. Single-Case Experimental Design. In: Madden GJ, Dube WV, Hackenberg TD, Hanley GP, Lattal KA, editors. *APA handbook of behavior analysis: Vol 1. Methods and principles*. Washington, DC: American Psychological Association; 2013:107–126.
8. Blampied NM. Single-Case Research Designs and the Scientist-Practitioner Ideal in Applied Psychology. In: Madden GJ, Dube WV, Hackenberg TD, Hanley GP, Lattal KA, editors. *APA handbook of behavior analysis: Vol 1. Methods and principles*. Washington, DC: American Psychological Association; 2013:177–197.
9. Gonzalez R. *Data analysis for experimental design*. New York, NY: Guilford Press; 2009.
10. Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference*. Boston, New York: Houghton Mifflin Company; 2002.
11. Woods B, Aguirre E, Spector AE, Orrell M. Cognitive stimulation to improve cognitive functioning in people with dementia. *Cochrane Database Syst Rev*. 2012;2:CD005562.
12. Dugard P, File P, Todman J. *Single-case and Small-n Designs: A Practical Guide to Randomization Tests*. New York, NY: Taylor & Francis Group; 2012.
13. Baer DM. Perhaps it would be better not to know everything. *J Appl Behav Anal*. 1977;10:167–172.
14. Morgan DL, Morgan RK. Single-participant research design. Bringing science to managed care. *Am Psychol*. 2001;56:119–127.
15. Backman CL, Harris SR, Chisholm JA, Monette AD. Single-subject research in rehabilitation: a review of studies using AB, withdrawal, multiple baseline, and alternating treatments designs. *Arch Phys Med Rehabil*. 1997;78:1145–1153.
16. Schwartz MF, Marin OS, Saffran EM. Dissociations of language function in dementia: a case study. *Brain Lang*. 1979;7:277–306.
17. Dechamps A, Fasotti L, Jungheim J, et al. Effects of different learning methods for instrumental activities of daily living in patients with Alzheimer's dementia: a pilot study. *Am J Alzheimers Dis Other Demen*. 2011;26:273–281.
18. Davies J, Howells K, Jones L. Evaluating innovative treatments in forensic mental health: A role for single case methodology? *J Forens Psychiatry Psychol*. 2007;18:353–367.
19. Barlow DH, Nock MK, Hersen M. *Single Case Experimental Designs: Strategies for Studying Behavior Change*. 3rd ed. Boston, MA: Allyn and Bacon; 1984.
20. Leitenberg H. The use of single-case methodology in psychotherapy research. *J Abnorm Psychol*. 1973;82:87–101.
21. Brenske S, Rudrud EH, Schulze KA, Rapp JT. Increasing activity attendance and engagement in individuals with dementia using descriptive prompts. *J Appl Behav Anal*. 2008;41:273–277.
22. Bourgeois MS. Effects of memory aids on the dyadic conversations of individuals with dementia. *J Appl Behav Anal*. 1993;26:77–87.
23. Arntzen E, Steingrimsdottir HS, Brogård-Antonsen A. Behavioral Studies of Dementia: Effects of Different Types of Matching-to-Sample Procedures. *Eur J Behav Anal*. 2013;40:17–29.
24. Heard K, Watson TS. Reducing wandering by persons with dementia using differential reinforcement. *J Appl Behav Anal*. 1999;32:381–384.
25. Baker JC, Hanley GP, Mathews RM. Staff-administered functional analysis and treatment of aggression by an elder with dementia. *J Appl Behav Anal*. 2006;39:469–474.
26. Altus DE, Engelman KK, Mathews RM. Using family-style meals to increase participation and communication in persons with dementia. *J Gerontol Nurs*. 2002;28(9):47–53.
27. Engelman KK, Altus DE, Mathews RM. Increasing engagement in daily activities by older adults with dementia. *J Appl Behav Anal*. 1999;32:107–110.
28. Watson PJ, Workman EA. The non-concurrent multiple baseline across-individuals design: an extension of the traditional multiple baseline design. *J Behav Ther Exp Psychiatry*. 1981;12(3):257–259.
29. Horner RD, Baer DM. Multiple-probe technique: a variation of the multiple baseline. *J Appl Behav Anal*. 1978;11(1):189–196.
30. Ulman JD, Sulzer-Azaroff B. Multielement baseline design in educational research. In: Ramp E, Semb G, editors. *Behavior analysis: Areas of research and application*. Enblewood Cliffs, NJ: Prentice Hall; 1975.
31. Runci S, Doyle C, Redman J. An empirical test of language-relevant interventions for dementia. *Int Psychogeriatr*. 1999;11:301–311.
32. Hartmann DP, Hall RV. The changing criterion design. *J Appl Behav Anal*. 1976;9:527–532.
33. Fitterling JM, Martin JE, Gramling S, Cole P, Milan MA. Behavioral management of exercise training in vascular headache patients: an investigation of exercise adherence and headache activity. *J Appl Behav Anal*. 1988;21:9–19.
34. Sidman M. *Tactics of scientific research: Evaluating experimental data in psychology*. New York: Basic Books; 1960.
35. Branch MN, Pennypacker HS. Generality and generalization of research findings. In: Madden GJ, Dube WV, Hackenberg TD, Hanley GP, Lattal KA, editors. *APA handbook of behavior analysis: Vol 1. Methods and principles*. Washington, DC: American Psychological Association; 2013:151–175.
36. Hussian RA. Modification of Behaviors in Dementia via Stimulus Manipulation. *Clin Gerontol*. 1988;8(1):37–43.
37. Nolan BA, Mathews RM, Harrison M. Using external memory aids to increase room finding by older adults with dementia. *Am J Alzheimers Dis Other Demen*. 2001;16:251–254.
38. Dwyer-Moore KJ, Dixon MR. Functional analysis and treatment of problem behavior of elderly adults in long-term care. *J Appl Behav Anal*. 2007;40(4):679–683.
39. Sellers DM. The Evaluation of an Animal Assisted Therapy Intervention for Elders with Dementia in Long-Term Care. *Activ Adapt Aging*. 2006;30(1):61–77.