the british psychological society
promoting excellence in psychology

British Journal of Health Psychology (2021)

www.wileyonlinelibrary.com

What makes implementation intentions (in)effective for physical activity among older adults?

Lisa Marie Warner[1]*, Lena Fleig[1], Julia Katharina Wolff[2,3], Jan Keller[4], Ralf Schwarzer[5], Samuel Nyman[6] and Susanne Wurm[3]

[1]Department of Psychology, MSB Medical School Berlin, Germany
[2]IGES Institute Berlin, Germany
[3]Department of Prevention Research and Social Medicine, Institute for Community Medicine, University Medicine Greifswald, Germany
[4]Health Psychology, Freie Universität Berlin, Germany
[5]SWPS University of Social Sciences and Humanities, Warsaw, Poland
[6]Bournemouth University Clinical Research Unit, Bournemouth University, UK

Objectives. For most populations, implementation intentions (IIs) facilitate physical activity (PA). However, for older adults, previous studies found mixed evidence for the effectiveness of this behaviour change technique. To examine which characteristics of IIs predict successful enactment, the content of older participants' IIs formed within a self-regulatory intervention to prompt PA was analysed.

Design. A sample of $N = 126$ German speaking adults aged 64 and older formed up to six IIs for PA and reported their enactment 5 weeks later.

Methods. Controlling for age and sex, multilevel models tested associations between characteristics of IIs (e.g., chronological rank of II, hetero- and homogeneity, specificity, presence of certain cues) and enactment.

Results. Significantly related to enactment were: the chronological rank of an II (first IIs superior to last IIs), greater heterogeneity in activities, greater specificity of when-cues, and greater use of pre-existing routines.

Conclusions. Participants were more likely to enact their IIs 5 weeks later if they planned different (heterogeneous) activities, created IIs with more specific when-cues (e.g., on Monday at 9 am), and in particular a routine (e.g., after breakfast). They also enacted the first three IIs (chronological rank of II) more often than the last three IIs. Future experimental studies should test whether providing instructions to create IIs based on the above significant characteristics lead to more effective health behaviour change among older adults.

**Correspondence should be addressed to Lisa Marie Warner, MSB Medical School Berlin, Rüdesheimer Straße 50, 14197 Berlin, Germany (email: lisa.warner@medicalschool-berlin.de).*

DOI:10.1111/bjhp.12563

Implementation intentions (IIs; Gollwitzer, 1993, 1999) were developed to bridge the so-called intention-behaviour gap, describing the phenomenon that forming good intentions is no guarantee for enacting them (Gollwitzer, 2014). IIs increase the likelihood of target behaviours by asking individuals to create 'if-then' plans that connect cues to actions ('If situation X arises, then I will perform the goal-directed response Y').

Compared to younger adults, older adults seem to translate their intentions into behaviour more easily (Hagger, Chatzisarantis, & Biddle, 2002). However, at least one third of Germans aged 60 years and older who intend to be as active as recommended, fail to reach recommended physical activity (PA) levels (Bauman, Merom, Bull, Buchner, & Fiatarone Singh, 2016; Krug et al., 2013). Reasons for this intention-behaviour gap among older adults can relate to difficulties with the initiation of activities (e.g., forgetting, distractions, temptations), losing track (e.g., old habits, competing goals), or getting exhausted while trying – psychological barriers that can be reduced by forming IIs (Gollwitzer, 2014).

Are IIs effective to promote physical activity among older adults?

Meta-analyses show that IIs increase the uptake of various health behaviours by medium to large effects (d = .61; Gollwitzer & Sheeran, 2006; Sheeran, 2002). If only interventions using IIs for PA are considered, the effect sizes are smaller (standardised mean differences of .25 to .31; Bélanger-Gravel, Godin, & Amireault, 2013; Silva, São-João, Brizon, Franco, & Mialhe, 2018). In complex interventions – including IIs along other behaviour change techniques (BCTs; Michie et al., 2013) – IIs did not always result in positive changes in PA (e.g., Bull et al., 2018; Finne et al., 2018; Michie, Abraham, Whittington, McAteer, & Gupta, 2009), but often belong to the active intervention ingredients (e.g., Grimmett et al., 2019; Howlett, Trivedi, Troop, & Chater, 2019; Olander et al., 2013; Webb, Joseph, Yardley, & Michie, 2010; Williams & French, 2011). However, among older adults, a meta-analysis (French, Olander, Chisholm, & Mc Sharry, 2014; Table 2) showed that ten BCTs had a negative impact on older adults' PA, including *action planning*, *planning of social support*, and *coping planning* (planning is similar to IIs; see Hagger et al., 2016).

Although we know that older adults prefer IIs including slower-paced physical activities in more frequent but shorter bouts (Alley, Schoeppe, Rebar, Hayman, & Vandelanotte, 2018), we need to understand whether certain IIs also increase the likelihood to enact PA. Therefore, this study aimed to (1) code systematically what older adults write down in IIs for PA, and (2) identify which characteristics of their IIs are associated with successful enactment.

Which II characteristics contribute to successful enactment?

In theory, IIs help close the intention-behaviour gap by forming a mental link between a situational cue and a behavioural response (Gollwitzer, 1993, 1999). IIs are assumed to increase the accuracy and speed of the detection of situational cues and to initiate less effortful (nearly automatic) responses (Gollwitzer, 1999). The encounter and detection of pre-formulated, contextual cues in real life is hence a prerequisite for behavioural initiation. As vague or ambiguous cues are more likely to be missed compared to precise specifications of opportunities to act (Gollwitzer, 1999; de Vet, Oenema, & Brug, 2011), cue detection and, thereby, successful plan enactment should increase with more detailed specificity of situational cues (Fleig et al., 2017).

Increasing the number of situational cues, either by formulating multiple IIs or by choosing a higher frequency of opportunities to act, such as 'daily' instead of 'Mondays' should increase the likelihood of cue encounter and enactment. However, increasing the number and with that possibly also the heterogeneity of IIs might come at the cost of weaker mental cue-behaviour links, a higher chance of forgetting or ignoring of cues, or interference between different IIs (Gollwitzer, 2014). Multiple similar IIs could possibly serve as repeated rehearsal that strengthens the mental cue-behaviour link (Gollwitzer, 1999). We investigate, whether forming several different IIs or similar IIs, which only differ in their cues but not in their behavioural response, is more beneficial for enactment. For this purpose, we define the concept of *heterogeneity* of a set of IIs as referring to how many different cues and/or activities individuals write into their IIs. We define the concept of *homogeneity* of a set of IIs as referring to the presence of multiple IIs that contain the same behavioural response.

What has been found to be effective about characteristics of IIs?

It has been recommended, '. . .that researchers examine how best to specify the if/then components of IIs' (Hagger et al., 2016, p. 824) and several field studies have revealed first results. For example, de Vet et al. (2011) rated characteristics of up to three IIs for PA among middle-aged working adults. They concluded that the mere number of IIs did not matter but that more *specific* IIs resulted in higher levels of overall PA. In their randomized controlled trial (RCT) with middle-aged working adults, Epton and Armitage (2017) used volitional help sheets containing various situational cues and physical activities to choose from in order to form IIs. They found that providing participants with more specific situational cues did not increase the effect of IIs on overall PA. However, a study by Fleig et al. (2017) that coded various characteristics of actions plans (similar to IIs, Hagger et al., 2016), found that more specific time-based cues were positively associated with the enactment of post-rehabilitation PA plans among working and retired middle- to older-aged individuals. Another study among middle- to old-aged couples found that plan enactment was more likely, if cues included a *routine* and not a time (e.g., 'after breakfast' instead of 'at 9 am'; Keller et al., 2017). This study also found that the chances of enactment for plans formed at the beginning of a planning sheet with up to five plans were higher than those formed towards the end. In the only other study that investigated this so-called *plan rank* (chronological rank of plans) among up to three plans, plank rank was not related to enactment (Fleig et al., 2017). Fleig et al. (2017) further found that leaving some flexibility in the activities planned was more effective than being too specific in formulating the behavioural response, whereas Keller et al.'s (2017) study could not replicate this finding. Evidence on the essence of effective IIs seems to vary across samples, and to date no study has focused on the characteristics of IIs for PA in older retired adults. Therefore, to develop new evidence on the essence of effective IIs with older adults, we coded the content of IIs created by older participants in an RCT testing a complex intervention with several motivational and volitional BCTs to increase PA.

Aims of the current study

First, we explore whether the *number* – and for the first time – the level of *heterogeneity* and *homogeneity* across a set of IIs were associated with enactment. We hypothesise that the *chronological rank of IIs* would be negatively associated with enactment, that is IIs made at the beginning of the worksheet would be enacted more likely than those made

towards the end. Second, we hypothesize that *specificity* ratings of all cues and the behavioural response would help enact IIs (i.e., the more detailed cues and behavioural responses, the more likely IIs are enacted). Finally, we hypothesize that using *routine-based* cues in IIs would be more predictive of enactment then *time-based* cues IIs.

Methods

Participants and procedure

The secondary data stem from a RCT testing a complex behaviour change intervention with motivational and volitional BCTs to increase PA in a group setting (without exercising, Appendix B lists all BCTs, Warner, Wolff, Ziegelmann, Schwarzer, & Wurm, 2016) against a parallel intervention for social volunteering and a passive control group (PREFER-II trial, funded by German Federal Ministry of Education and Research, 01ET1001B). In total, $N = 647$ community-dwelling adults aged 64 and older were assessed for eligibility: exercising less than twice a week for 30 min and not acutely physically impaired or disabled. Of these, $n = 360$ adults were randomized to the three groups. To investigate characteristics of the PA-related IIs, this study used the II worksheets created in the intervention group for PA ($n = 126$), of which $n = 115$ (91%) completed the worksheets plus the 5-week follow-up diary on II enactment. As participants actively approached the institute after having read newspaper articles or advertisements for the project 'Active Retirement', the majority was motivated to be more active (Table 2; Appendix D).

Approval was obtained from the ethics commission of the German Psychological Society (DGPs-SW 02_2012). The trial was preregistered at www.clinicaltrials.gov (NCT01577134). More information on the RCT can be found in Warner et al. (2016) and Warner, Wolff, Spuling, & Wurm (2019).

Design

Study design

The RCT consisted of five measurement points (Warner et al., 2016). For present analyses, data on II characteristics stem from the II-worksheet completed during the intervention session (i.e., T1; May-September, 2012). Data on the enactment of IIs stem from the PA diary 5 weeks post-intervention (T2).

Implementation intention worksheet

The worksheet to formulate IIs was one out of ten BCTs implemented as part of a single face-to-face group intervention session with three to eight participants lasting approximately 2 hr and 40 min, including two short breaks (more information in Appendices B and C and Warner et al., 2016; Wolff, Warner, Ziegelmann, & Wurm, 2014). After having reviewed the positive consequences of PA, selecting activities, watching a role model video and setting their own activity goals, the interventionists explained that 'if-then sentences' can help translate goals into actions. Participants also learned that choosing good cues to action and linking these to specific physical activities is important for effective 'if-then sentences'. Participants then received several II examples displayed on slides and verbally explained (e.g., 'If I have finished breakfast, then I do my gymnastic exercises', see Appendix C for more). Participants were told that cues to action

can consist of situations (e.g., exact times, routines), objects (e.g., yoga mat, sports bag), or persons (e.g., friends, family). After these instructions, a worksheet was distributed to every participant to write down up to six individual IIs prompting the BCT, 'Action planning (including implementation intentions)' (BCT [1.4] according to Michie et al., 2013) but not 'Problem solving/coping planning' (BCT [1.2] Michie et al., 2013; Appendix C - translated worksheet). Two facilitators (postdoctoral researchers) were available for questions during the completion of the worksheets, but did not proactively support or correct the development of IIs.

Measures

Enactment of IIs (dependent variable)

To assess whether participants enacted their IIs in daily life, we used data from the 10-day activity diary completed at 5-weeks following the intervention. In end-of-day assessments, participants reported which physical activities they had performed. Matched to participants' II, raters coded how often the planned activity appeared in the diary. This resulted in up to six enactment sum scores related to respective IIs (e.g., if the first II contained cycling and cycling was reported five times in the diary, the enactment score for this II was five). By counting only activities that participants had included in their IIs during the intervention, the dependent variable represents the conditional outcome of II enactment, but not general activity levels after the intervention (Sniehotta, 2009). This enactment of PA-related IIs reflects a more proximal outcome than overall PA and is suitable to examine the effects of plan characteristics (Fleig et al., 2017; Keller et al., 2017).

Implementation intentions (independent variables)

First, participants' IIs were screened for completion and adherence to instructions. IIs were coded as *invalid* if no or invalid information was entered in the 'if'- and/or 'then'-part (e.g., for invalid 'if'-entry: 'If my friend would not be that lazy, . . .'; Appendix A defines valid components). Characteristics (see below) were only coded for IIs that were previously coded as complete and valid.

This study focussed on four different components of IIs as independent variables: the occasion cue (i.e., when-cue), location cue (i.e., where-cue), and the social cue (i.e., who-cue) from the 'if-part', as well as the planned behavioural response (i.e., which PA) from the 'then-part'. The coding manual for this study was based on previous coding manuals (Fleig et al., 2017; Keller et al., 2017), and can be downloaded here. Two independent and trained raters scored each II separately. Cohen's kappa inter-rater agreement before consensus discussions ranged from fair (i.e., .38 for categorisation of when-cue characteristics) to almost perfect (i.e., .91 for specificity of social cues; Landis & Koch, 1977). Consensus scores for final variables were derived by raters' discussions in the presence of the first author.

Characteristics of a set of IIs

The *number* of created IIs was established by counting each II that contained a valid if- and then-component to create a score ranging from 0 to 6.

The *chronological rank* of *II* represents the order in which participants formed their IIs (from 0 = first to 5 = sixth).

The *heterogeneity* of when-, where- and social cues as well as behavioural responses were assessed from 1 to 6, depending on *how many different cues and/or activities* participants wrote into their IIs, divided by the total number of valid IIs created.

As a measure for *homogeneity*, the number of IIs concerning the *same behavioural response (i.e., then-part), that is physical activity,* was assessed from 0 to 5, divided by the total number of valid IIs created.

Characteristics of specific parts of IIs

Specificity of IIs

Specificity was coded from 1 (unspecific/vague) to 3 (highly specific) for each component (when-cue, where-cue, social cue, behavioural response) of an II. The specificity was coded as follows: unspecific, if either no information was inserted, or if the entry described no concrete occasion/location/social partner(s); moderately specific, if at least some or ambiguous information about the occasion/location/social partner(s) was provided; highly specific, if the information was very clear and unambiguously usable as a cue to action (Appendix A provides details and examples).

Type of when-cues: Times, routines, opportunities

For when-cues, two types of cues, namely the *presence of time points* (e.g., 'Monday 8 a.m.') *and* the *presence of routines* (e.g., 'after breakfast'), were coded as two separate dichotomous variables with '1 = present' and '0 = not present'. In addition, the *number of opportunities per week* was extracted by coding the frequency of weekly occasions in an II (e.g., seven for 'daily').

Data analyses

In all models, continuous predictors were grand-mean-centred (except for chronological rank of II; coded from 0 to 5) and dichotomous predictors included '0' as reference category. Missing data were treated using the full information maximum-likelihood (FIML) procedure for all analyses in Mplus8 (Muthén & Muthén, 2017; amount of missing data in Appendix F).

Model A aimed at investigating whether the characteristics of a set of up to six IIs predicts enactment. The number of IIs as well as ratings of heterogeneity and homogeneity of IIs were entered as predictors for enactment in a regression with Poisson distribution (to account for count data of enactment) with maximum likelihood estimator, controlling for participants' age and sex (0 = female, 1 = male). As age and the number of IIs were significant predictors of enactment, these two variables were modelled as covariates in subsequent models. To enable an interpretation of the magnitude of effects, significant B coefficients were transferred to incidence rate ratios (IRRs; Hilbe, 2011).

Further, two multilevel models (Model B and Model C) with characteristics of specific parts of IIs (Level 1) nested within participants (Level 2) predicted enactment as Level 1 outcome, using a Poisson prediction as well as a maximum-likelihood estimation with Monte Carlo integration. Random-intercept models showed a small intraclass correlation

coefficient of *ICC* = 0.03 indicating that most variation in IIs was observed at Level 1 (still multilevel modelling was used to account for the nested data structure, Huang, 2018).

Model B tested whether the Level 1 predictors of chronological rank of IIs, specificity levels of cues and behavioural responses predict enactment. To allow for between-person (Level 2) differences in the effects of Level 1 predictions, all Level 1 predictors were modelled as random effects (Barr, Levy, Scheepers, & Tily, 2013). Model C focused on testing certain characteristics of when-cues as predictors of plan enactment, by only examining the first two *situational* IIs on the worksheet (because few participants included information on IIs 3–6 regarding objects and persons). Three different cue characteristics were modelled as Level 1 predictors of enactment: The presence of time points, presence of routines, and number of opportunities per week (modelled as random effects).

Results

Descriptive results

Descriptive results for the sample and the quantity of IIs

The total sample comprised N = 126 participants, of which 73% were women, 63% were high school graduates and 46% lived together with a partner. At baseline, participants' age was M = 70.53 years (SD = 5.01, 64–91 years). Of these 126 individuals, eight participants (6%) did not write anything into their worksheet and 15 (12%) created if-then-sentences that were coded as invalid II (e.g., 'If appointment, then I go there'). The remaining n = 103 participants with at least one valid II created on average M = 2.94 (SD = 1.94, range 1–6) IIs. A total of n = 93 participants with valid IIs returned their activity diaries at the 5-weeks follow-up (see Appendices D & E for comparisons between the n = 93 and the n = 126 sample in, e.g., health status, cognitive abilities, reported acceptability of the intervention). On average, facilitators rated fidelity as high, participants were satisfied with the intervention, and satisfaction scores between the analysed and overall sample did not differ (see Appendix E). Across the 10-day diary, participants enacted a mean of 2.22 (SD = 3.08; range: 0–18) II-related activities.

Descriptive results

Highest *heterogeneity* in IIs was found with regard to different activities planned (M = 2.67, SD = 1.32, range 1–6). Regarding when-cues, participants specified M = 2.24 different occasions (SD = 1.10, range 0–4). With respect to social cues, M = 0.67 (SD = 0.79, range 0–3) different activity partners were planned. Where-cues were specified with an average of M = 0.39 different cues (SD = 0.76, range 0–4). The number of IIs concerning the same PA (homogeneity) was on average M = 0.72 (SD = 1.08, range 0–3).

To summarise descriptive statistics of the *specificity* of IIs, scores for up to six IIs were averaged. Participants were most specific in forming IIs on the planned PA (M = 2.26, SD = 0.63, range 1–3), followed by when-cues (M = 1.74, SD = 0.48, range 1–3), while social cues (M = 1.23, SD = 0.32, range 1–2), and where-cues (M = 1.13, SD = 0.25, range 1–2.20) were rather unspecific.

Across all participants, 67% used routines and 27% used time points in at least one II. 53% of participants specified a number of opportunities to be active per week in any of

Table 1. Poisson regression model with number of IIs, heterogeneity indicators, and homogeneity of the behavioural response as predictors of enactment

Model A	*B*	*SE*	*p*	95% CI	
Intercept	−1.75	0.74	.017	−3.196	−0.308
Age	.03	0.01	.001	0.012	0.048
Sex	.12	0.11	.275	−0.096	0.339
Number of formed IIs	.25	0.04	<.001	0.172	0.323
Heterogeneity of when-cues	−.01	0.18	.937	−0.369	0.340
Heterogeneity of where-cues	−.38	0.25	.129	−0.861	0.109
Heterogeneity of behavioural responses	.72	0.19	<.001	0.348	1.084
Homogeneity of behavioural responses: Number of times same behaviour was planned	−.01	0.15	.950	−0.305	0.286

Note. n = 93 participants; CI = confidence interval.

their IIs. On average, participants planned a mean of 3.75 (SD = 5.24, range 0-28) opportunities per week to be physically active.

Prediction of enactment by number, heterogeneity and homogeneity across IIs (Model A)

The number of formed IIs as well as their heterogeneity and homogeneity were tested as predictors of enactment. As Table 1 shows, more IIs were associated with a higher likelihood of enactment. With each additional II, participants enacted an additional 28% of physical activities (IRR = 1.28, 95% CI [1.184, 1.378]; all other predictors held constant). Furthermore, the heterogeneity of the behavioural responses significantly predicted enactment. The IRR showed that planning one more activity was associated with twice (IRR = 2.05, 95% CI [1.294, 2.799]) as many enacted physical activities (all other predictors held constant). None of the other heterogeneity indicators, nor homogeneity predicted enactment.

The heterogeneity of when-cues correlated at −.83 with the heterogeneity of social cues, causing problems with multicollinearity. As when-cues were more frequent than social cues, we excluded the heterogeneity of social cues (including or excluding the heterogeneity of when-cues resulted in the same significant predictors). Regarding covariates, older age positively predicted enactment (one more year predicting a 3% increase in enactment, IRR = 1.03, 95% CI [1.012, 1.049]), whereas sex was unrelated.

Prediction of enactment by chronological rank and specificity of IIs (Model B)

To test associations of the chronological rank of an II (Hypothesis 1) and II specificity (Hypothesis 2) with enactment, Model B (Table 2) used multilevel modelling controlling for age and number of IIs. As expected, the chronological rank of IIs was negatively related to enactment. The IRR (0.79, 95% CI [0.659, 0.917]) indicated that for each II further down the worksheet, enactment decreased by 21% (all other predictors held constant). Descriptives showed that the first three IIs were often enacted: first M = 2.58 (SD = 3.23) enactments per 10 days, second M = 3.84 (SD = 3.33), and third M = 2.28 (SD = 2.81). Whereas, the fourth (M = 1.18, SD = 2.54), fifth (M = 0.62, SD = 0.99), and sixth (M = 1.36, SD = 2.11) were less enacted (however, only 11 participants formed six IIs). Regarding specificity indicators, only when-cue specificity was significantly and positively

Table 2. Poisson multilevel model with chronological rank of IIs and specificity indicators as predictors of enactment

	Fixed effects				Random effects			
			95% CI				95% CI	
Model B	*B* (*SE*)	*p*	*LL*	*UL*	Variance (*SE*)	*p*	*LL*	*UL*
Intercept	.55 (0.21)	.007	0.148	0.956	0.24 (0.19)	.205	−0.133	0.621
Age	.03 (0.03)	.207	−0.017	0.080	–	–	–	–
Number of formed IIs	.01 (0.08)	.851	−0.136	0.165	–	–	–	–
Chronological rank of II	−.24 (0.08)	.004	−0.402	−0.074	0.07 (0.05)	.178	−0.030	0.160
Specificity of when-cues	.48 (0.22)	.027	0.054	0.914	1.23 (0.61)	.044	0.033	2.426
Specificity of where-cues	−.26 (0.22)	.242	−0.691	0.175	0.43 (0.38)	.253	−0.307	1.164
Specificity of social cues	−.82 (0.45)	.066	−1.697	0.054	1.00 (0.83)	.229	−0.631	2.639
Specificity of behavioural response	−.23 (0.19)	.225	−0.606	0.143	1.26 (0.55)	.023	0.178	2.350

Note. $n = 93$ participants with 230 observations; CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

related to enactment, whereas specificity of where-cues, social cues, and behavioural responses were not significant. Each unit increase in the rating of when-cue specificity predicted an increase of 62% enactments (IRR 1.62, 95% CI [0.925, 2.321], all other predictors held constant).

Prediction of enactment by characteristics of 'when-cue' (Model C)

To test whether including routines as opposed to certain times or the number of opportunities is associated with II enactment (Hypothesis 3), they were entered as predictors in Model C controlling for age and number of IIs. Table 3 shows that only the presence of routine-based cues was associated with enactment, whereas the presence of time points or the number of planned opportunities per week were unrelated to enactment. Including a routine-based cue was associated with an 80% increase in likelihood of enactment (IRR = 1.80, 95% CI [1.036, 2.560]).

Conclusions

In this study, secondary data were derived by coding the content of IIs older adults created in a complex behaviour change PA RCT (Warner et al., 2016). Enactment of IIs was regressed on codings of number, ranking, hetero- and homogeneity of IIs, specificity and presence of certain when-cues in three multilevel models with ascending resolution (from characteristics of a set of IIs to specificity ratings of cues to the presence of routines, times, and number of opportunities). Participants were more likely to enact their IIs 5 weeks later, if they planned different (heterogeneous) activities, used more specific when-cues (e.g., 'Monday 6 am'), and included a routine (e.g., 'after breakfast'). They enacted the first three IIs on the worksheet (chronological rank) more often than the last three IIs.

Table 3. Poisson multilevel models with 'when-cue' characteristics as predictors of enactment

	Fixed effects				Random effects			
			95% CI				95% CI	
Model C	*B* (*SE*)	*p*	*LL*	*UL*	Variance (*SE*)	*p*	*LL*	*UL*
Intercept	.40 (0.17)	.018	0.067	0.728	0.44 (0.18)	.017	0.079	0.790
Age	.04 (0.02)	.109	−0.009	0.086	–	–	–	–
Number of formed IIs	.06 (0.08)	.427	−0.094	0.222	–	–	–	–
Presence of routines in when-cues	.58 (0.21)	.006	0.169	0.984	0.15 (0.23)	.508	−0.299	0.604
Presence of time in when-cues	−.07 (0.37)	.856	−0.790	0.656	0.90 (0.66)	.171	−0.387	2.181
Number of opportunities per week in when-cues	−.01 (0.04)	.831	−0.083	0.067	0.01 (0.01)	.346	−0.010	0.029

Note. n = 82 participants with 121 observations; CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

Characteristics of a set of IIs

Each additional PA planned (heterogeneity of then-part) was associated with twice as many enactments, whereas planning the same activity (homogeneity) at different occasions and locations (heterogeneity of cues) did not affect enactment. Our results may give support for the notion, that more heterogeneity in planned behaviours – but not cues – increases the likelihood of enactment. However, using heterogeneous when-cues and where-cues was not detrimental to enactment. This means that for promoting PA among older people, providing instructions to plan different physical activities could be more effective than providing instructions to think of different occasions and locations to be active. A higher number of IIs did not seem to interfere or overburden participants, but was associated with a higher likelihood of enactment in Model A. This 'quantity effect' was however not replicated in Models B and C. Therefore, we refrain from interpreting it, as well as the age effect that only emerged in Model A. To our knowledge, previous studies have not addressed heterogeneity and homogeneity of IIs. Future experimental research could test whether homogeneity of cues is more effective than novel and heterogeneous cues. The longer learning and planning history of older adults might enable them to transfer previously successful cues to actions to other behavioural domains.

Characteristics of specific parts of IIs

The chronological rank of IIs predicted enactment: the first three IIs – especially the second one – were implemented more often. Previous research also indicates that plan rank matters when individuals have the opportunity to formulate up to five plans (Keller et al., 2017), but not so in interventions with up to three plans (Fleig et al., 2017; de Vet et al., 2011). Formulating up to six IIs can be challenging. It is likely, that participants first specified those IIs that they were most committed to and activities and occasions that they anticipated to fit best to their daily routines.

Our finding that more specific when-cues were associated with higher rates of enactment is in line with previous research (Fleig et al., 2017; de Vet et al., 2011). It also integrates well with theory that IIs, which contain specific when-cues, have a high chance to be detected upon encounter (Gollwitzer, 2014). The results of the current study, however, suggest that besides specificity of cues, the type of the occasion cue makes a difference for enactment. Routines rather than exact times or the number of opportunities were found to be the key to successful IIs for PA among older adults. This supports previous research among older (Fleig et al., 2017) and middle-aged and younger adults (Keller et al., 2017). In our study, it was found to be beneficial if older adults envisioned to enact their PA either *while* engaging in their usual routine (e.g., 'If I brush my teeth') or after a routine (e.g., 'If I have finished breakfast'). In the context of dietary planning across the whole lifespan, especially morning routines predicted healthy eating (Domke, Keller, Fleig, Knoll, & Schwarzer, 2019), which would align well with adults' preferences to be active in the morning (Alley et al., 2018). Although experimental research among younger adults suggests that time-based or routine-based cues are similarly effective for habit formation (Keller et al., 2021), routines require less active 'cue-monitoring' than times. Maybe this explains why performance in event-based prospective memory tasks is less affected by age then performance in time-based tasks in laboratory settings (however not in naturalistic settings; Henry, MacLeod, Phillips, & Crawford, 2004).

Possible problems in planning interventions for physical activity among older adults

Previous research found that IIs help older adults to buffer effects of declining prospective memory on delayed task performance in other life domains (e.g., Brom & Kliegel, 2014). However, including IIs in complex interventions was negatively related to PA outcomes for older adults (French et al., 2014). According to French et al. (2014, French, Banafa, Williams, Taylor, Brown, 2020), this may be due to two reasons. First, planning might be *cognitively difficult* for older adults. Secondly, older adults may have a *lower need for planning* and prefer flexible weekly activities.

To assess whether formulating up to 6 IIs might have been cognitively challenging, we investigated how many valid IIs were made: Out of 126 participants, 118 worked on the worksheets, but only 103 generated at least one valid II. For those 15 participants without at least one valid II – who did not cognitively differ from the group with valid worksheets (Appendix D) – closer supervision might have been useful. Planning can be experienced as burdensome (Fleig et al., 2016), and older adults benefit more from interviewer-assisted planning while younger adults benefit more from self-administered planning (Ziegelmann, Lippke, & Schwarzer, 2006). The fact that the vast majority (88%) of our participants created valid plans and that they did not differ in cognitive abilities from those who did not create plans (see Appendix D), is, however, reason to believe that most older adults, who volunteer to participate in a social-cognitive PA intervention, are cognitively able to use this BCT with minimal assistance.

The interventionists approached eight participants with individualised prompts, because they were not working on their worksheet. Seven openly expressed refusal to work with the II worksheet. This rejection might have been linked with a *lower need to plan*. It has been suggested that retirees have more leisure time and more flexible schedules (French et al., 2014, 2020) and/or less conflicting goals due to fewer obligations (Riediger, Freund, & Baltes, 2005). Some studies also report that retirees highly value flexibility, prefer internal/mood cues (e.g., 'if I feel energetic') and exercise habitually (French et al., 2020). The lack of time is, however, still amongst the most prominent barriers for PA during retirement (Moschny, Platen, Klaassen-Mielke, Trampisch, & Hinrichs, 2011). Our findings do not stand against planning interventions for older adults per se, but for the use of established routines rather than exact time points to promote PA. Participants who openly expressed reactance toward IIs were the minority. Reactance should still be assessed in future studies (e.g., Ungar, Sieverding, Schweizer, & Stadnitski, 2015) along with potential moderators (e.g., personality, history of planning biography, previous experiences with planning; Engel & Kuhl, 2015; French et al., 2020).

Strengths & limitations

A strength of our study is that we assessed enactment around 5 weeks after IIs were generated and not immediately like experimental research on IIs. Also, the enactment of IIs was operationalized as the number of reported enactments of pre-planned physical activities – a measure theoretically most closely related to IIs (as opposed to overall PA). This indirect measurement probably faces a lower risk of social desirability, rather than directly asking whether IIs were fulfilled. And by coding enactment per II, we accounted for the multilevel structure of IIs nested within participants.

However, it needs to be noted that we performed *post-hoc* secondary analyses of a subsample of an RCT to increase physical activity that was not effective (Warner et al., 2016), which bears different risks of bias (detailed in Appendix G). As most participants formed between one and three IIs (Appendix F), the number of observations used in the

multilevel models did not reach the theoretically possible number of observations. The FIML procedure was, however, only used for single missing values (due to minor coding problems with some II characteristics) and did not apply to non-existent IIs nor missing enactment values. The study was powered for the main trial, however, not for these sub-analyses. The sample consisted of self-selected community-dwelling older adults. It was hence biased towards healthier, cognitively fitter, more motivated, and higher educated individuals than the German population aged 65+ years (details see Appendix D; and Warner et al., 2019). Also, some inter-rater reliabilities were poor, limiting the reliability of our results and highlighting the need for more training for coders.

Future research

In field studies, participants usually receive written or spoken instructions and then formulate IIs according to their understanding of these instructions with or without supervision. To investigate possible cognitive challenges and moderators for the preference or reluctance to use IIs as a BCT among older adults, future studies could use the think-aloud method (French, Cooke, McLean, Williams, & Sutton, 2007). Upon identification of key characteristics of IIs, or more elaborated formulations of planning instructions, RCTs with factorial designs testing supposedly effective mechanisms of IIs against one another should be conducted (e.g., Keller et al., 2021 for a factorial design in younger adults).

One participant of our study wrote onto the worksheet for IIs 'I want to be active when I feel like it!'. This tendency to focus on short-term mood cues for PA, such as making best use of one's energy rather than long-term health effects, is supported by socio-emotional selectivity theory and warrants future research (Carstensen, 1992; French et al., 2020; Pimm et al., 2016).

Implications

IIs are low-cost and easy to include in behaviour change interventions at a population level. However, this is only useful if individuals self-generate *effective* IIs. This is particularly important for older adults for whom the current evidence-base is mixed. Bearing our low number of participants and the similarity in participant characteristics (mostly female, highly educated, and motivated, see Appendices D–G) in mind, we conclude that planning instructions for older adults might highlight the importance of creating IIs with very specific information about the occasion for several different physical activities and to embed their intended activities into existing daily routines. These suggestions however need to be tested experimentally in larger samples.

Acknowledgements

The authors would like to thank the German Centre of Gerontology and a team of highly motivated student assistants for helping to conduct the study. The PREFER II study (Personal Resources of Elderly People With Multimorbidity: Fortification of Effective Health Behaviour) was funded by the German Federal Ministry of Education and Research (grant number 01ET1001B). Ethical consent was granted from the ethics commission of the German Psychological Society (#DGPs-SW 02_2012). Open access funding enabled and organized by ProjektDEAL.

Conflicts of interest

All authors declare no conflict of interest.

Author contribution

Lisa Marie Warner: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **Lena Fleig:** Conceptualization (equal); Data curation (equal); Methodology (equal); Writing – review & editing (equal). **Julia Katharina Wolff:** Conceptualization (equal); Data curation (equal); Investigation (equal); Methodology (equal); Supervision (equal); Validation (equal); Writing – review & editing (equal). **Jan Keller:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Validation (equal); Writing – review & editing (equal). **Ralf Schwarzer:** Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal). **Samuel Nyman:** Writing – review & editing (equal). **Susanne Wurm:** Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal).

Data availability statement

Anonymised and edited data of the PREFER study (with German variables and value labels only) is available free of charge from the FDZ-DZA as Scientific Use File in SPSS- and Stata-format. For reasons of data protection, signing a data distribution contract is required to receive the data sets. If you want to use PREFER data for secondary analysis, please contact the research data centre of the German Centre of Gerontology (https://www.dza.de/en/research/fdz/prefer).

References

Alley, S. J., Schoeppe, S., Rebar, A. L., Hayman, M., & Vandelanotte, C. (2018). Age differences in physical activity intentions and implementation intention preferences. *Journal of Behavioral Medicine*, *41*, 406–415. https://doi.org/10.1007/s10865-017-9899-y

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bauman, A., Merom, D., Bull, F. C., Buchner, D. M., & Fiatarone Singh, M. A. (2016). Updating the evidence for physical activity: Summative reviews of the epidemiological evidence, prevalence, and interventions to promote 'active aging'. *The Gerontologist*, *56*, 268–280. https://doi.org/10.1093/geront/gnw031

Bélanger-Gravel, A., Godin, G., & Amireault, S. (2013). A meta-analytic review of the effect of implementation intentions on physical activity. *Health Psychology Review*, 7(1), 23–54. https://doi.org/10.1080/17437199.2011.560095

Brom, S. S., & Kliegel, M. (2014). Improving everyday prospective memory performance in older adults: Comparing cognitive process and strategy training. *Psychology and Aging*, *29*, 744–755. https://doi.org/10.1037/a0037181

Bull, E. R., McCleary, N., Li, X., Dombrowski, S. U., Dusseldorp, E., & Johnston, M. (2018). Interventions to promote healthy eating, physical activity and smoking in low-income groups: A systematic review with meta-analysis of behavior change techniques and delivery/context.

International Journal of Behavioral Medicine, *25*, 605–616. https://doi.org/10.1007/s12529-018-9734-z

Carstensen, L. L. (1992). Social and emotional patterns in adulthood: Support for socioemotional selectivity theory. *Psychology and Aging*, 7, 331–338. https://doi.org/10.1037//0882-7974.7.3.331

de Vet, E., Oenema, A., & Brug, J. (2011). More or better: Do the number and specificity of implementation intentions matter in increasing physical activity? *Psychology of Sport and Exercise*, *12*, 471–477. https://doi.org/10.1016/j.psychsport.2011.02.008

Domke, A., Keller, J., Fleig, L., Knoll, N., & Schwarzer, R. (2019). What makes a good action plan? Characteristics and enactment of fruit and vegetable plans. *Appetite*, *142*, 104351. https://doi.org/10.1016/j.appet.2019.104351

Engel, A. M., & Kuhl, J. (2015). Personality and planning: The interplay between linear and holistic processing. In M. D. Mumford & M. Frese (Eds.), *The psychology of planning in organizations: Research and applications* (pp. 58–88). Routledge/Taylor & Francis Group.

Epton, T., & Armitage, C. J. (2017). Does situation-specificity affect the operation of implementation intentions? *Behavior Therapy*, *48*, 860–869. https://doi.org/10.1016/j.beth.2017.08.003

Finne, E., Glausch, M., Exner, A. K., Sauzet, O., Stölzel, F., & Seidel, N. (2018). Behavior change techniques for increasing physical activity in cancer survivors: A systematic review and meta-analysis of randomized controlled trials. *Cancer Management and Research*, *10*, 5125–5143. https://doi.org/10.2147/CMAR.S170064

Fleig, L., Gardner, B., Keller, J., Lippke, S., Pomp, S., & Wiedemann, A. U. (2017). What contributes to action plan enactment? Examining characteristics of physical activity plans. *British Journal of Health Psychology*, *22*, 940–957. https://doi.org/10.1111/bjhp.12263

Fleig, L., McAllister, M. M., Chen, P., Iverson, J., Milne, K., McKay, H. A., . . . Ashe, M. C. (2016). Health behaviour change theory meets falls prevention: Feasibility of a habit-based balance and strength exercise intervention for older adults. *Psychology of Sport and Exercise*, *22*, 114–122. https://doi.org/10.1016/j.psychsport.2015.07.002

French, D. P., Banafa, R., Williams, S., Taylor, C., & Brown, L. J. (2020). How does the understanding, experience, and enactment of self-regulation behaviour change techniques vary with age? A thematic analysis. *Applied Psychology: Health and Well-Being*, *13*(1), 239–260. https://doi.org/10.1111/aphw.12243

French, D. P., Cooke, R., McLean, N., Williams, M., & Sutton, S. (2007). What do people think about when they answer theory of planned behaviour questionnaires? *Journal of Health Psychology*, *12*, 672–687. https://doi.org/10.1177/1359105307078174

French, D. P., Olander, E., Chisholm, A., & Mc Sharry, J. (2014). Which behaviour change techniques are most effective at increasing older adults' self-efficacy and physical activity behaviour? A systematic review. *Annals of Behavioral Medicine*, *48*, 225–234. https://doi.org/10.1007/s12160-014-9593-z

Gollwitzer, P. M. (1993). Goal achievement: The role of intentions. *European Review of Social Psychology*, *4*(1), 141–185. https://doi.org/10.1080/14792779343000059

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503. https://doi.org/10.1037/0003-066X.54.7.493

Gollwitzer, P. M. (2014). Weakness of the will: Is a quick fix possible? *Motivation and Emotion*, *38*, 305–322. https://doi.org/10.1007/s11031-014-9416-3

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, *38*, 69–119. https://doi.org/10.1016/S0065-2601(06)38002-1

Grimmett, C., Corbett, T., Brunet, J., Shepherd, J., Pinto, B. M., May, C. R., & Foster, C. (2019). Systematic review and meta-analysis of maintenance of physical activity behaviour change in cancer survivors. *International Journal of Behavioral Nutrition and Physical Activity*, *16*(1), 1–20. https://doi.org/10.1186/s12966-019-0787-4

Hagger, M., Chatzisarantis, N., & Biddle, S. (2002). A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: Predictive validity and the contribution of

additional variables. *Journal of Sport & Exercise Psychology*, *24*(1), 3–32. https://doi.org/10.1123/jsep.24.1.3

Hagger, M. S., Luszczynska, A., de Wit, J., Benyamini, Y., Burkert, S., Chamberland, P.-E., … Gollwitzer, P. M. (2016). Implementation intention and planning interventions in health psychology: Recommendations from the Synergy Expert Group for research and practice. *Psychology & Health*, *31*, 814–839. https://doi.org/10.1080/08870446.2016.1146719

Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and Aging*, *19*(1), 27–39. https://doi.org/10.1037/0882-7974.19.1.27

Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge, UK: Cambridge University Press.

Howlett, N., Trivedi, D., Troop, N. A., & Chater, A. M. (2019). Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. *Translational Behavioral Medicine*, *9*, 147–157. https://doi.org/10.1093/tbm/iby010

Huang, F. L. (2018). Multilevel modeling myths. *School Psychology Quarterly*, *33*, 492–499. https://doi.org/10.1037/spq0000272

Keller, J., Fleig, L., Hohl, D. H., Wiedemann, A. U., Burkert, S., Luszczynska, A., & Knoll, N. (2017). Which characteristics of planning matter? Individual and dyadic physical activity plans and their effects on plan enactment. *Social Science & Medicine*, *189*, 53–62. https://doi.org/10.1016/j.socscimed.2017.07.025

Keller, J., Kwasnicka, D., Klaiber, P., Sichert, L., Lally, P., & Fleig, L. (2021). Habit formation following routine-based versus time-based cue planning: A randomized controlled trial. *British Journal of Health Psychology*, *26*, 807–824. https://doi.org/10.1111/bjhp.12504

Krug, S., Jordan, S., Mensink, G. B. M., Müters, S., Finger, J., & Lampert, T. (2013). Physical activity: Results of the German Health Interview and Examination Survey for Adults (DEGS1). *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, *56*, 765–771. https://doi.org/10.1007/s00103-012-1661-6

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychology*, *28*, 690–701. https://doi.org/10.1037/a0016136

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., … Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, *46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Moschny, A., Platen, P., Klaassen-Mielke, R., Trampisch, U., & Hinrichs, T. (2011). Barriers to physical activity in older adults in Germany: A cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity*, *8*(1), 121. https://doi.org/10.1186/1479-5868-8-121

Muthén, L. K., & Muthén, B. O. (1998– 2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Olander, E. K., Fletcher, H., Williams, S., Atkinson, L., Turner, A., & French, D. P. (2013). What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *10*(1), 1–15. https://doi.org/10.1186/1479-5868-10-29

Pimm, R., Vandelanotte, C., Rhodes, R. E., Short, C., Duncan, M. J., & Rebar, A. L. (2016). Cue consistency associated with physical activity automaticity and behavior. *Behavioral Medicine*, *42*, 248–253. https://doi.org/10.1080/08964289.2015.1017549

Riediger, M., Freund, A. M., & Baltes, P. B. (2005). Managing life through personal goals: Intergoal facilitation and intensity of goal pursuit in younger and older adulthood. *The Journals of*

Gerontology Series B: Psychological Sciences and Social Sciences, *60*, P84–P91. https://doi.org/10.1093/geronb/60.2.p84

Sheeran, P. (2002). Intention – Behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, *12*(1), 1–36. https://doi.org/10.1080/14792772143000003

Silva, M. V. A., São-João, T. M., Brizon, V. C., Franco, D. H., & Mialhe, F. L. (2018). Impact of implementation intentions on physical activity practice in adults: A systematic review and meta-analysis of randomized clinical trials. *PLoS One*, *13*, e0206294. https://doi.org/10.1371/journal.pone.0206294

Sniehotta, F. F. (2009). Towards a theory of intentional behaviour change: Plans, planning, and self-regulation. *Britisch Jounral of Health Psychology*, *14*, 261–273. https://doi.org/10.1348/135910708x389042

Ungar, N., Sieverding, M., Schweizer, F., & Stadnitski, T. (2015). Intervention-elicited reactance and its implications: Let me eat what I want. *Zeitschrift Für Psychologie*, *223*, 247–256. https://doi.org/10.1027/2151-2604/a000226

Warner, L. M., Wolff, J. K., Spuling, S. M., & Wurm, S. (2019). Perceived somatic and affective barriers for self-efficacy and physical activity. *Journal of Health Psychology*, *24*, 1850–1862. https://doi.org/10.1177/1359105317705979

Warner, L. M., Wolff, J. K., Ziegelmann, J. P., Schwarzer, R., & Wurm, S. (2016). Revisiting self-regulatory techniques to promote physical activity in older adults: Null-findings from a randomised controlled trial. *Psychology & Health*, *31*, 1145–1165. https://doi.org/10.1080/08870446.2016.1185523

Webb, T., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet Research*, *12*(1), e1376. https://doi.org/10.2196/jmir.1376

Williams, S. L., & French, D. P. (2011). What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour – And are they the same? *Health Education Research*, *26*, 308–322. https://doi.org/10.1093/her/cyr005

Wolff, J. K., Warner, L. M., Ziegelmann, J. P., & Wurm, S. (2014). What do targeting positive views on ageing add to a physical activity intervention in older adults? Results from a randomised controlled trial. *Psychology & Health*, *29*, 915–932. https://doi.org/10.1080/08870446.2014.896464

Ziegelmann, J. P., Lippke, S., & Schwarzer, R. (2006). Adoption and maintenance of physical activity: Planning interventions in young, middle-aged, and older adults. *Psychology & Health*, *21*, 145–163. https://doi.org/10.1080/14768320500188915

Received 25 March 2021; revised version received 15 September 2021

Supporting Information

The following supporting information may be found in the online edition of the article:

Appendix A Coding manual for implementation intentions.
Appendix B Context and content of the intervention.
Appendix C Instructions for participants on how to create implementation intentions.
Appendix D Description of study design and participants.
Appendix E Evaluation by facilitators and participants.
Appendix F Descriptive statistics for Models A, B, C.
Appendix G Risk of Bias.