**Supplementary items**

Appendix 1. Database Search Terms

|  |  |
| --- | --- |
| Line | Term |
| 1 | elder\*.ti,ab. |
| 2 | older.ti,ab. |
| 3 | aged.ti,ab |
| 4 | aging.ti,ab. |
| 5 | ageing.ti,ab. |
| 6 | geriatric.ti,ab. |
| 7 | senior.ti,ab. |
| 8 | retire\*.ti,ab. |
| 9 | veteran.ti,ab. |
| 10 | pensioner\*.ti,ab. |
| 11 | "old age\*".ti,ab. |
| 12 | "over 60\*".ti,ab. |
| 13 | "over 65\*".ti,ab. |
| 14 | sixties.ti,ab. |
| 15 | “over 70”.ti,ab. |
| 16 | seventies.ti.ab. |
| 17 | "over 80".ti,ab. |
| 18 | eighties.ti,ab. |
| 19 | "care home\*".ti,ab |
| 20 | “care facilit\*”.ti,ab. |
| 21 | "nursing home\*".ti,ab. |
| 22 | "residential home\*".ti,ab. |
| 23 | "residential facilit\*".ti,ab. |
| 24 | institution\*.ti.ab. |
| 25 | "community care".ti,ab. |
| 26 | "assisted living".ti,ab. |
| 27 | "green house project".ti,ab. |
| 28 | "sheltered housing".ti,ab. |
| 29 | dement\*.ti,ab. |
| 30 | Alzheimer\*.ti,ab. |
| 31 | Aged/ |
| 32 | Aging/ |
| 33 | Aged 80 and over/ |
| 34 | Frail Elderly/ |
| 35 | Middle Aged/ |
| 36 | Homebound Persons/ |
| 37 | Veterans/ |
| 38 | Geriatrics/ |
| 39 | Retirement/ |
| 40 | Alzheimer Disease/ |
| 41 | Dementia/ |
| 42 | Assisted Living Facilities/ |
| 43 | Homes for the Aged/ |
| 44 | Housing for the Elderly/ |
| 45 | Group Homes/ |
| 46 | Senior Centers/ |
| 47 | Long-Term Care/ |
| 48 | Institutionalization/ |
| 49 | Nursing Homes/ |
| 50 | 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 |
| 51 | (indoor\* adj2 garden\*).ti,ab. |
| 52 | (interior\* adj2 garden\*).ti,ab. |
| 53 | (internal adj2 garden\*).ti,ab. |
| 54 | (garden\* adj2 room).ti,ab. |
| 55 | "enclosed garden\*".ti,ab. |
| 56 | (indoor\* adj2 horticultur\*).ti,ab. |
| 57 | (interior adj2 horticultur\*).ti,ab. |
| 58 | (internal adj2 horticultur\*).ti,ab. |
| 59 | (indoor\* adj2 plant\*).ti,ab. |
| 60 | (interior adj2 plant\*).ti,ab. |
| 61 | houseplant\*.ti,ab. |
| 62 | “house plant\*”.ti,ab. |
| 63 | "potted plant\*".ti,ab. |
| 64 | aquari\*.ti,ab. |
| 65 | "fish tank\*".ti,ab. |
| 66 | (natur\* adj2 indoor\*).ti,ab. |
| 67 | (natur\* adj2 inside).ti,ab. |
| 68 | (natur\* adj2 home\*).ti,ab. |
| 69 | (natur\* adj2 image\*).ti,ab. |
| 70 | (natur\* adj2 scene\*).ti,ab. |
| 71 | (natur\* adj2 pictur\*).ti,ab. |
| 72 | (natur\* adj2 photo\*).ti,ab. |
| 73 | (natur\* adj2 art\*).ti,ab. |
| 74 | (natur\* adj2 mural\*).ti,ab. |
| 75 | (natur\* adj2 depict\*).ti,ab. |
| 76 | (natur\* adj2 represent\*).ti,ab. |
| 77 | (natur\* adj2 paint\*).ti,ab. |
| 78 | (natur\* adj2 film\*).ti,ab. |
| 79 | (natur\* adj2 movie\*).ti,ab. |
| 80 | (natur\* adj2 video\*).ti,ab. |
| 81 | (natur\* adj2 recording\*).ti,ab. |
| 82 | (natur\* adj2 sound\*).ti,ab. |
| 83 | (natur\* adj2 audio\*).ti,ab. |
| 84 | (natur\* adj2 virtual\*).ti.ab |
| 85 | (natur\* adj2 simulat\*).ti.ab |
| 86 | (natur\* adj2 artificial\*).ti.ab |
| 87 | 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 |
| 88 | 50 and 87 |

The search was made in the MEDLINE database via Ovid host, with Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE and Versions(R) selected. Each search line was submitted independently into the basic search field, with title and abstract fields selected (except for MeSH terms, which used a /, as shown in steps 31-49). The search lines relating to older adults/residential settings and indoor nature were combined with OR at steps 50 and 87 respectively. At step 88, these were combined with AND. The results were filtered for English language studies. There were no date restrictions on the search.

Appendix 2. Example of completed data extraction form

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Reviewer initials: NY | Today’s date: 05/07/17 | | | Unique study number: 5 | |
| Article title: Effects of watching aquariums on elders' stress | | | | | Year published:  1990 |
| 1st author: DeSchriver | | Publication type: Original article | | | |
| Publication title: Anthrozoos | | | | | |
| URL: http://www.tandfonline.com/doi/pdf/10.2752/089279391787057396 | | | | | |
| Study Question(s)/Aim(s)/Objective(s)/Hypotheses: To determine whether viewing real fish or videos of fish can produce a relaxational effect or reduce physiological stress in elders. | | | | | |
| Design: Three groups RCT | | | Setting: publicly subsidised housing unit for elderly persons in Maryland, US | | |
| Sample size: 27  Age range (average): not stated  Sex: 21 females (78%), 6 males (22%)  Ethnicity: not stated  Detail any other relevant pre-intervention measures –for example gardening experience, health status, co-morbidities: None stated | | | Recruitment procedure:  Not stated | | |
| Random assignment to groups (Y/N/can’t tell/not applicable): Yes | | |
| Unit of allocation to groups: at individual level | | |
| Number of groups:  Group 1  Exposure/intervention(s): Fish aquarium  No. participants: 9  Age range (average): median 75y  Sex: 8 female (89%), 1 male (11%)  Ethnicity: not stated  Group 2  Exposure/interventions(s): Fish video  No. participants: 9  Age range (average): median 73y  Sex: 7 females (78%), 2 males (22%)  Ethnicity: not stated  Group 3  Exposure/interventions(s): placebo video tapes (static)  No. participants: 9  Age range (average): median 76y  Sex: 6 females (67%), 3 males (33%)  Ethnicity: not stated | | | Details of nature element(s):  The fish aquarium was a landscaped ten-gallon tank stocked with nine fish (two black mollies, two red wag swordtails, two gold wag moons, two pineapple swordtails, and one catfish).  The fish videotape was a modified version of a videotape produced by the Candle Corporation that showed (almost exclusively in close-up shots)  a variety of colorful tropical fish swimming in an aquarium. The audio portion of the videotape was enhanced by superimposing the sound of a stream  trickling over rocks.  Study participants were tested individually  in the arts-and-crafts room of the apartment complex where they resided | | |
| Frequency and duration of intervention(s):  8-minute session, once per week for 3 weeks | | |
| Whole study duration: 3 weeks | | |
| Year(s) study took place: not stated | | |
| Measures/instruments/tools   * Cognitive stressor – reading aloud an emotionally stimulating article for 3 min * Treatment Evaluation Scale constructed using a revised version of Beard and Ragheb’s (1980) shortened Leisure Satisfaction Scale (LSS) * Cardiovascular activity as a proxy for physiological stress. Composed of pulse rate, skin temperature and musculoskeletal tension | Statistical methods  Test for between-groups differences in the pre-post test score changes.  The test(s) used is not stated but appears to be ANOVA. | | | Key results/effect sizes/findings  TES mean scores:  Fish aquarium: 12.33  Fish videotape: 12.89  Control: 12.23  Between-groups diff NS (F [2, 24]=0.08, p=0.93.  Pre & post-test mean scores  Pulse rate (bpm):  Aquarium: 72.02 to 73.19  Fish video: 71.50 to 71.32  Control: 72.39 to 72.61  *F [2, 24]=0.50, p=0.24*  Skin temp (degrees Fahrenheit):  Aquarium: 94.08 to 94.23  Fish video: 91.16 to 91.81  Control: 91.92 to 91.98  *F [2, 24]=0.50, p=0.30*  Muscle tension (mV):  Aquarium: 23.46 to 21.40  Fish video: 26.24 to 24.04  Control video: 26.26 to 24.62  *F [2, 24]=0.37, p=0.35* | |
| When were Measures/instruments/tools administered?  Data were collected during all three sessions but only data collected during the third session was included in the analysis. The author stated that the first two sessions were necessary to ‘acclimatise’ participants to the conditions and avoid habituation.  TES: post-test (after 3rd session)  Physiological measures: pre-test (means of 2 readings taken 1 min apart) and post-test (means of 8 readings taken 1 min apart) |
| Author-identified strengths  The cognitive stressor wasn’t used in the first two (acclimatising) sessions to control for possible habituation.  No other activities were undertaken in the testing room during the study | Author-identified limitations  Very small sample sizes | | | Author-identified conclusions  They suggested the fish videotape has a greater impact on reducing stress than the aquarium. **But they don’t** - the F scores showed no difference in effect across the 3 conditions for both the TES and the physiological measures. | |
| Reviewer-identified strengths | Reviewer-identified limitations  There were no analyses for within-group score changes on the physiological outcomes.  The first line of the discussion says that aquarium viewers had a **decrease** in pulse rate. Table 1 shows it was an **increase**.  Participants were heard “conversing with others between sessions about their favourite fish” – potential contamination?  Extremely limited demographic information and no consideration of how demographics might confound  No information on reliability or validity of the TES  Test-retest reliability for the EMG was poor (0.56)  Only the measures analysed in the final week were taken for analysis. | | | Reviewer-identified conclusions  All three conditions led to similar tiny increases in skin temperature and slightly larger decreases in muscle tension. Change in pulse rate was variable.  But there were no significant differences between the groups for any of the physiological variables.    Given the small sample size and lack of within-group analyses, it would be unreasonable to draw any conclusions about the ability of viewing a fish aquarium or fish videotape to induce relaxation/reduce stress in this population. | |
| Funding source: not stated | | | | | |
| Conflicts of interest: not mentioned by the authors but the fish videotape was produced by the Candle corporation. This corporation also made a donation of “software equipment” to the group. Given the authors’ unsupported conclusions re the videotape working best, there may be some CoI here. | | | | | |

Appendix 3. Descriptions of individual studies

Aslakson (2010) used a **RCT** to compare the ability of music therapy (their intervention, my comparator) and **nature videos** (their control, my intervention) to improve **agitation, engagement** and **functional behaviour (social interaction, task performance, problem solving)** in 40 nursing home residents in the Midwestern US. Three sessions of music therapy or nature videos, each 30-40 mins in duration, were delivered during one week. Pre and post-test measures were collected using the **Wisconsin Agitation Inventory (WAI)**, **Functional Behaviour Profile (FBP)** and an **Engagement** variable developed by the researcher. ANOVA and ANCOVA were used to assess between-group differences. Both groups experienced a significant decrease in agitation (music: 44.12 to 30.18; nature 42.64 to 33.82; F=5.83, p<0.05) but there were no significant between-groups differences. There was a significant between-groups difference in pre to post engagement scores (music: 3.3 to 4.5; nature: 3.4 to 2.3; F=24.54, p<0.01). There was a significant between-groups difference in the social interactions component of the FBS (F=0.93, p<0.05) but there were no descriptive statistics reported for this measure, so exact score changes, or direction of changes could not be determined. Conclusion: watching a nature video or taking part in music therapy sessions significantly improved agitation in people with dementia. Nature videos were associated with significantly decreased engagement. There was no information on the specific content of the nature video.

Barnicle and Midden (2003) and Midden and Barnicle (2004) used a **controlled clinical trial** to evaluate the effects of a **horticulture activity programme** on **psychological wellbeing (affect)** in 62 residents of two long-term care facilities in Missouri, US. The intervention group lived in facility 1, and received 1x 1hr guided horticulture session per week for 7 weeks. The control group lived in facility 2, received usual care, and was told the horticulture programme would start in 7 weeks (waiting list). The **Affect Balance Scale** toolwas administered at pre and post-test. One-way ANOVA was used to assess within-group changes. Two-way ANOVA was used to test for between-group differences. The intervention group had a pre to post-test increase in mean ABS from 5.42 to 7.61 (F=3.17, p=0.08). The control group had a pre to post-test decline in mean ABS from 4.29 to 3.00 (F=0.70, p=0.40). The pre to post-test change in ABS score was significantly greater in the intervention than the control group (F=6.78, p=0.01). Conclusion: residents who took part in guided horticultural activities had a near-significant increase in psychological wellbeing over a 7-week programme.

Brown *et al* (2004) used a **cluster RCT** to examine the effects of **indoor gardening** on **socialization, activities of daily living (ADLs),** and **perceptions of loneliness** in 66 elderly nursing home residents of two facilities in the south-eastern US. The intervention group lived in facility 1 and received 1x indoor gardening session (duration not stated) per week for 5 weeks. The control group lived in facility 2 and received 20 minute visits from facility staff. The **UCLA Loneliness Scale (V3), Revised Social Provisions Scale** and **Minimum Data Set Physical Functioning Scale for ADLs** were administered at pre and post-test. There were significant within-group improvements for both groups in loneliness and several socialisation measures, but there were no significant between-group differences. The intervention group had significant improvements in three ADLs associated with upper body movement, but the equivalent results for the control group were not reported. In a second phase, the control group received a 2-week indoor gardening programme. This group’s post-test scores did not significantly differ from their visits scores, and there were no significant within-group changes for this group during phase II. The 5-week gardening programme was significantly favourable over the 2-week programme on four socialisation subscales and two physical functioning scales. However, it was unclear which scores (post-test scores, pre-to-post change) were used in this latter analysis. Conclusions: A 5-week gardening programme was no more beneficial than 20 minute visits for improving several psychological and physical health outcomes in elderly nursing home residents. Some measures indicated that a gardening programme conducted once per week for 5 weeks was more beneficial to health than a programme conducted twice per week for 2 weeks.

Chung *et al* (2016) used **one-group** **pre/post design** to explore effects of **media presentations containing nature** on **agitated behaviours** of 23 dementia patients residing in a long-term care facility in Salt Lake City, US. Participants viewed nature DVDs for 7-10 mins, once per day, 3x per week for 2-4 weeks. Agitated behaviours through **daily nursing records**. MANOVA was used to compare these against baseline records taken 2 weeks prior to the study. Frequency of one behaviour type, ‘hitting a resident’ changed to a near-significant level during the intervention, but the direction and size of the change was not reported. Conclusion: in 6 out of 7 measures, short nature DVDs had no effect on agitated behaviours in dementia patients. The effect on ‘hitting a resident’ was unclear.

Cohen-Mansfield and Werner (1998) employed a **quasi-experimental crossover** design to assess the effects of two **simulated enhanced environments** on **behaviour, mood, and pacing/wandering** of 27 nursing home residents who pace frequently. The home’s location was not specified. Following a 2-week baseline data collection period, ‘nature’ and ‘home’ scenes were erected in separate corridors. The scenes were rotated, and were in place for a total of 4 weeks over a period of 6 weeks. The **Confusion Inventory**, **Lawton’s Modified Behaviour Stream** (mood) and **Cohen-Mansfield Agitation** **Inventory**, as well as direct researcher observations of participants’ **location, body position, pacing/wandering, exit-seeking and trespassing** behaviours were employed during scene and no-scene periods. Location and exit-seeking, trespassing, and pacing/wandering were also recorded with **automatic counting** and **ambulatory tracking devices**, respectively. Baseline/no-scene vs scene scores were compared using t-tests and Wilcoxon alpha. When either enhancement was in situ, participants exhibited (non-significantly) less lying down, exit-seeking and trespassing. During the nature enhancement, participants additionally exhibited less sleeping, and during the home enhancement, they additionally exhibited less pacing and less standing. A significant increase in the mood sub pleasure was observed between no-scene and nature (1.05 v 1.11, p<0.05), but not the home scenes. For participants who displayed agitated behaviours during baseline, large reductions in frequency of agitation episodes occurred in three of four agitation categories for nature, and two of four categories for home. No between-groups statistical comparisons were made for any measures. Conclusion: A nature scene enhancement was associated with fewer instances of agitated behaviour, less time sleeping and increased pleasure compared with home scenes, but several other behaviours, including exit-seeking and trespassing, were similar for both enhancements. Pacing/wandering only decreased in the home enhancement.

Collins and O’Callaghan (2008) used a **one-group pre/post** design to study the effects of a **horticulture activity programme** on **psychological wellbeing** in 18 residents of a low-income assisted living facility in Nevada, US. Participants attended 2-hour interactive horticulture classes once per week for 4 weeks. The quantitative part of the study employed **Pearlin & Schooler’s Mastery Scale** and **self-reported health** and **self-reported happiness** were each scored with 1 item. Measures were taken pre (t1) and post-test (t2), and additionally at 5-month follow-up (t3) for a subset of participants. Paired t-tests were used to analyse differences in scores between each of the time points. There were significant improvements in personal mastery from t1-t2 (*t*= -6.75, p=0.001), t1-t3 (*t*= -4.07, p=0.005), and near significant improvement from t2-t3 (*t=* -2.0, p=0.086). There were significant improvements in self-reported health from t1-t2 (*t*= -4.12, p=0.001), t1-t3 (*t*= -3.99, p=0.005), and t2-t3 (*t=* -0.75, p=0.02). There were significant improvements in self-rated happiness from t1-t2 (*t*= 2.2, p=0.042), t1-t3 (*t*= -2.65, p=0.033), and near-significant improvements from t2-t3 (*t=* -2.05, p=0.08). Conclusions: A short-term horticulture intervention including taking personal responsibility for plants seemed to be beneficial for participants in terms of three wellbeing and quality of life measures: perception of personal mastery, self-reported health and self-rated happiness. However, the lack of control group means effects are not certainly attributable to the intervention.

DeSchriver and Riddick (1990) used a **three-group** **RCT** to examine the ability of viewing real fish in an **aquarium tank**, watching **videos of fish**, or watching videos of static to induce **relaxation** and reduce **physiological stress** in 27 elders living in a publicly subsidised housing unit in Maryland, US. Participants attended sessions for 8 minutes once per week for 3 weeks. Cardiovascular activity was taken as a proxy for physiological stress, including **pulse rate, skin temperature** and **muscle tension** and was measured before and after each session. Only results from the third session were analysed. The statistical test was not stated, but between-group post-test comparisons were made for each measure, controlled for pre-test scores: there were no significant differences. There were also no significant within-group changes in any measures from pre to post test. Conclusion: Neither a fish aquarium or fish videotape was able to reduce physiological stress levels when viewed in 1x 8-minute session.

D’Andrea *et al* (2008) used a **RCT** to determine whether a **horticulture therapy programme** could slow **cognitive decline** in a sample of 40 nursing home residents with Alzheimer’s disease living in a dementia special care unit in New York, US. Half the participants were randomly assigned to receive 30-45 minutes of indoor horticultural activities twice per week for 12 weeks. The other half served as controls, and continued regular scheduled activities which included music sessions, reminiscence discussion and socialising. **Minimum Data Set Plus (MDS+)** was administered ‘throughout the study’ and **Test for Severe Impairment (TSI)** was employed pre and post-test. No data for the MDS+ was reported, but the authors indicated improvements in functional level, mood, behaviour, ADLs and wellbeing among all participants in the horticulture group. TSI scores slightly improved from pre to post-test in the horticulture group (19.5 vs 20.4) but worsened in the controls (21.5 vs 19.5). Because the baseline scores significantly differed, an independent t test was conducted on the difference scores (generated by subtracting the post-test scores from the pre-test scores). The results showed a significant difference between the groups (t=5.7, p<0.0005). Conclusion: Taking part in a 12 week HT programme was associated with maintenance of cognition in people with AZD.

Edwards and Beck (2002) used a **controlled quasi-experiment** to study the influence of **fish aquariums** on **nutritional intake** and **body weight** in 62 elderly individuals with Alzheimer’s Disease living in three dementia-specific units in Indiana, US. An aquarium was installed in the dining room of facilities 1 and 2 for eight weeks. A scenic ocean picture was installed in the dining room of facility 3 (control) for two weeks. Following a washout & further baseline period, facility 3 switched to the aquarium condition for eight weeks (waiting list). Participants were exposed to the installations during mealtimes (3x per day). **Body weight** was measured once per month, starting from 3 months prior to the start of the intervention. Nutritional Intake was operationalised as the **weight of food consumed** (in grams) at mealtimes and was measured at every meal (3x per day) during baseline and for the first 2 weeks of the aquarium*,* and then once weekly for the remaining 6 weeks. Paired t-tests were used to assess within-group changes in mean nutritional intake between baseline and treatment. In the intervention group nutritional intake: 1) increased 21.1% from baseline to the end of the two-week period in which measurements were taken daily (all facilities combined t score -7.276, p<0.001) and 2) increased 27.1% between baseline and the end of the 6 weeks in which measurements were taken weekly (t score -7.932, p<0.001). There were no significant changes in nutritional intake within the control group. The statistical test used to analyse change in weight was not stated. There was a significant increase in intervention group weight in the month the aquarium was introduced (M: 0.54 Ibs) and over the four months since introducing the aquarium (1.65 Ibs, p<0.001) No information was given about weight for the control group. There were no between-group comparisons for either measure. Conclusion: introducing an aquarium into the dining room of a dementia facility may lead to increased nutritional intake and body weight, but the lack of comparison with the control group means this is not certain.

Edwards and Beck (2013) used a **one-group pre/post** design to assess the influence of viewing an **aquarium** at mealtimes on **food intake** and **body weight** of 70 elderly individuals with dementia living in three specialised dementia units in North Carolina, US. An aquarium was introduced into the facility dining room for 8 weeks and was visible to all participants during mealtimes. Body weightwas measured once per month, beginning 3 months prior to the start of the intervention and continuing for 3 months afterwards. Food intakewas operationalised as the **weight of food consumed** (in grams) at mealtimes and was measured at every meal (3x per day) during the baseline period (phase I)*,* once per day during the first 2 weeks (phase II) of the intervention, and once per week for the remaining 6 weeks (phase III). Mealtime scores were averaged to produce mean daily food intake, and then daily intakes were averaged across each phase. Mean increase in food intake was 121.6 g between phases I and II (a significant increase) and 75.3 g (a non-significant increase) between phases II and III, making a total mean increase of 196.9g (a 25% increase) over the 10-week period. Repeated measures ANOVA found a significant main effect of phase on total food intake (F=85.7, p<0.001). Paired t tests with Bonferroni correction were employed to examine changes in body weight. A significant increase of 2.2 pounds (t=7.5, P=0.000 <0.05/3) occurred between the mean baseline weight (158.4 pounds) and the mean weight at the end of phase III (160.6 pounds). Conclusion: introducing an aquarium into the dining room of a dementia facility may lead to increased food intake and body weight, but the lack of control group means effects are not certainly attributable to the intervention.

Edwards *et al* (2014) used a **one-group pre/post** designto examine influence of an **aquarium** on **behaviour** and **psychological symptoms** of 71 people with dementia living in specialised dementia units in North Carolina and Florida, US. Following a 2-week baseline data collection period, an aquarium was placed in the activity room to be visible at mealtimes for all residents over the course of 8 weeks. An adapted version of the **Nursing Home Disruptive Behaviour Scale** was employed pre and post-test. Repeated measures mixed-model ANCOVA was used to examine pre to post-test changes. Overall behaviour scores significantly improved (mean 67.2 vs 58.2, F=15.6, p<0.001). Changes were significant along 4 behaviour domains: uncooperative (F= 4.76, p=0.033), irrational (F= 9.29, p=0.003), sleep (F=4.62, p=0.035) and inappropriate (F12.36, p=0.001), nearing significance for annoying behaviour (F= 3.81, p=0.055) and non-significant for dangerous behaviour. Conclusion: aquariums placed in a central location were associated with a significant decrease in problematic behaviours overall, and in four of six behavioural domains. A decreasing trend was observed in one additional behavioural category. The lack of control group means effects are not certainly attributable to the intervention.

Eggert *et al* (2015) used a **two-group quasi within-subject** design to assess whether viewing **preferred nature images** or **preferred music** impacts on **cognitive ability**, **engagement** and **dementia-related disordered behaviours** of 13 residents of a memory care unit. The location was not specified. Participants firstly selected one nature image from a selection representing themes from Appleton’s Prospect Refuge Theory, or one song from a selection of genres. They then looked at and talked about the image, or listened to and sang along with the song, with the researcher, in 90 minute one-to-one sessions, once per week for 4 weeks. The **Individualized Dementia Engagement and Activities Scale (IDEAS)** and **Cohen-Mansfield Agitation Inventory (CMAI)** were used at the start and end of each session. **The Montreal Cognitive Assessment (MoCA)** was used one week before and one week after the intervention period. Descriptive analysis was undertaken. For the MoCA clock-drawing, one participant’s score improved from 1 to 2. There were no improvements in the music group. For the MoCA memory recall, three of nine participants (33%) scored 2 points or more before the nature intervention compared with four of eight (50%) afterwards. Zero of four scored 2 points or more before the music intervention, compared with two of six (33.4%) afterwards. On the IDEAS, there were very small increases of between 0.7-1.8 points during each session for both interventions. Only two items from the CMAI, relating to verbally aggressive behaviours, were reported, and patterns were inconsistent across sessions for both interventions. Conclusion: The small sample size and mixed results meant the evidence from this study was insufficient to suggest that preferred nature images or music can affect agitation, engagement and cognition in persons with dementia.

Goto *et al* (2014) used a **controlled quasi-experiment** to determine the relative effects of a Snoezelen room and **indoor Japanese garden** on **behaviour** and **physiological stress levels** of 36 nursing home residents with Alzheimer’s and other forms of dementia living at a nursing home in New Jersey, US. Participants entered the environments for 15 minutes twice per week for 3 weeks (Snoezelen) or 4 weeks (garden). The interventions were separated by one year. Six participants experienced both environments. Physiological stress was operationalised as **pulse rate** and was taken during each session and post-session in the participant’s own rooms. **The Behavioural Assessment Checklist** was taken during each session. T-tests, Chi2 and one-way ANOVA were used to test effects of the interventions. Significantly more participants remained awake in the garden than the Snoezelen room (mean 12.63 vs 2.5, t=11.18, p<0.001). Significantly more participants moved around in the garden than the Snoezelen room (X2=50.44, p<0.001). Fewer participants chose to leave the garden than the Snoezelen room. The percentage of verbalisations in the garden room was higher (56%) than the Snoezelen room (24%). Affect and attention were not compared statistically, but more of the garden participants showed positive affect (bright, smile), whereas level of attention paid to the setting was generally higher in the Snoezelen room – notably the researchers observed that most attention was paid to a nature projection. There was an almost continuous decrease in average pulse rate in the garden of approximately 0.15-0.2 beats per minute. By contrast, pulse rate in the Snoezelen room was more variable, with an overall average increase of 0.06 beats per minute. Comparing the last 6 minutes in the garden with post-test measures taken in participant’s own rooms, garden pulse rates were significantly lower (p=0.034). There was no difference for the Snoezelen group in this respect (p=0.34). Regression to the mean may have been involved but this pattern did occur in the six participants who experienced both conditions. The study included a control group by taking 15 participants back to the garden location after it was dismantled, but no quantitative data for this group was reported. Conclusion: being in a Japanese garden appeared to somewhar reduce physiological stress (pulse rate), increase willingness to participate and improve verbalisations compared with a Snoezelen room.

Kieffer (2014) used a **cross-sectional** design to determine whether **images containing representational elements of nature** (REN) could increase perceived **self-reported wellbeing** in 20 residents of a senior independent living community in Minnesota, US. Participants viewed 4 pairs of photographs of senior living facilities public lounges. In each pair, one photo contained REN (water, fire, botanical motifs, natural materials) and the other did not. The participants answered a **questionnaire designed by the researcher** which ascertained how they thought they would feel if they were in the environments depicted. Perceived wellbeing was on five scales: refreshed – exhausting, distracted – attentive, relaxed – harried, irritable – patient and comfortable – uneasy. Descriptive analysis was undertaken. Pictures containing water were perceived as better for wellbeing than pictures without water on all five scales. Pictures containing natural materials were perceived as worse for wellbeing than pictures without natural materials on 4/5 scales. Results for fire and botanical motifs were mixed. Conclusions: Images depicting care home interiors containing water were perceived as better for wellbeing than images of interiors without water.

Kiyota (2009) used a **three group cluster RCT** to examine whether active or passive interaction with **indoor plants** would impact on **perceived restoration**, **helplessness** and **depression** in 30 residents of an elderly care facility in Canada. Plants were placed in the living room of six houses for 6 weeks. The plants were cared for by staff in three of the houses (passive group) and the residents in the other three (active group). Three further houses did not have plants and served as the control group. The active group participants also received 5 minutes of **horticulture tutoring** per week from the researcher. The **Modified Perceived Restoration Scale** and **Geriatric Depression Scale** were administered during baseline and once per week during the intervention. A mixed ANOVA was used with timepoint (x7) and group (x3) as factors. There were no significant group or time effects on perceived restoration, but a significant group x time interaction occurred (*F* 2.115, p=0.023). No post-hoc tests were performed but the mean PRS scores for the active group increased to a greater extent than the passive group. There were no significant effects of group, time or group x time on GDS. However, there was a trend of reducing scores over time for both the active and passive groups. Conclusion: Exposure to indoor plants had a significant positive effect on perceived restoration, and a positive trend in improving depression. Having active involvement in caring for plants was more beneficial than simply observing them.

Lee and Kim (2008) used a **one-group pre/post design** to determine the effect of an **indoor gardening programme** on **sleep, cognition** and **agitation** in 23 dementia patients living in an institution (location not specified). Participants took part for 1 hour twice per day every day for 28 days, and could also access their plants whenever they wanted. The **Modified Cohen-Mansfield Agitation Inventory** and 24-hour sleep diaries were collected every day during week 1 (baseline) and week 5 (final gardening week). The **Hasegawa Dementia Rating Scale - Revised** was employed once in week 1 and once in week 5. Paired t-tests were used to analyse pre/post-test changes. Significant pre to post-test improvements were observed in wake after sleep onset duration (75.22 vs 54.65 mins, *t* 2.781, p=0.011) and frequency (6.08 vs 2.21 occasions, *t* 3.568, p= 0.002), nap duration (158.43 vs 85.87 mins, *t* 7.933, p= <0.001) and frequency (3.18 vs 1.95 occasions, *t* 6.480, p=<0.001), Nocturnal Sleep Time (440.48 vs 483.52 mins, *t* -3.493, p=0.002), Nocturnal sleep efficacy (85.09 vs 89.62%, *t* -3.048 p=0.006), agitation (5.09 vs 3.13, *t* -4.002, p=0.001) and cognition (13.70 vs 17.48, *t* 12.044, p<0.001). Conclusion: participating in an indoor gardening programme was associated with improved sleep and cognition and reduced agitation in dementia patients. The lack of control group means effects are not certainly attributable to the intervention.

Martin (2011) used a **cluster-randomised crossover experiment** to determine whether **viewing landscape photographs** could reduce **agitation**, and whether there would be a dose-response effect in 22 nursing home residents with Alzheimer’s and a history of agitation in New York, US. Each wing was allocated to display landscape photos (experimental) or photos depicting the interior of their nursing home facility (control). Slide shows containing 180 photos were projected on a loop in a public area of the wing for five days between 9am and 5pm. Following a washout period of 9 days, the conditions were switched. **The Brief Agitation Ratings Scale** was employed at baseline, during the viewing periods and during the washout period. **Dose** was operationalised as the number of minutes the participant was within the perimeter of the display, regardless of whether they were looking at it. Two-sample t-tests or Mann Whitney U tests were performed to assess effects of treatment, time period, treatment by time period and first-order carryover. The group which saw the experimental display first was denoted EC. The group which saw the control display first was denoted CE. There was an overall downward trend in scores (i.e. overall agitation improved) in both groups throughout the course of the study. There was no significant treatment effect (*t=*0.97, p=0.34), period effect (*t*=1.96, p=0.07) or carryover effect (*t=*-0.68, p=0.50). There was a significant time by period interaction (*t*=2.52, p=0.02), but this was lost with a supplemental calculation which adjusted for scores from the periods immediately before the experimental display periods (*U=*45, P=0.33). After dropping outliers, agitation decreased as dose increased, and this was significant when fitted with a linear model (*R2*=0.15, *F*=4.45, p=0.05). However, as with the main analysis, significance was lost when adjusting for scores from the periods immediately before the experimental display periods. Odds of improvement in the experimental condition were approximately 12x greater for females than males (OR=12.00, p=0.04) and 5x greater for mild-moderately agitated participants than moderately-highly agitated participants (OR=4.67, p=0.09). Conclusion: there were no significant improvements in agitation between the nature and control condition.

Masuya *et al* (2014) used a **controlled pre/post quasi experiment** to determine the effect of a **horticultural activities programme** on **cognitive, psychological and physical functioning and QoL** of 18 elderly nursing home residents. The home’s location was not specified. The intervention group (n=9) received 30-40 minutes of horticulture activity once per week for six weeks. The control group (n=9) were recruited from a different floor of the nursing home, and received routine care. **The Vitality Index (VI), Geriatric Depression Scale (GDS-15), Activities of Daily Living Scale (ADL-20), Mini Mental State Exam (MMSE)** and a **Visual Analogue Scale for QoL** were employed pre and post-test. Two-way ANOVA with Bonferroni correction was used for within and between-groups testing. GDS-15 score improved in the intervention group from pre to post-test (5.7 vs 3.6, F=14.01, p<0.05), and a time x group interaction occurred (F=8.12, p<0.05). The QoL item ‘satisfaction with life’ improved in the intervention group from pre to post-test (75.6 vs 91.1, F=28.00, p<0.05) and a time x group interaction occurred (F=16.46, p<0.05). No significant pre/post-test changes occurred in any other measure for the intervention group. No significant pre/post-test changes were seen in any of the measures for the control group. No other between-group differences or interactions were found. Conclusions: Participation in a horticultural activities programme significantly improved short-term depression and satisfaction with life scores in a small sample of elderly nursing home residents.

Moyle *et al* (2018) used a **one-group pre/post** pilot study to explore the effects of a **Virtual Reality** **forest** on **emotions, apathy** and **engagement** of people with dementia living in two residential care facilities in Australia. The intervention consisted of a large immersive wall-mounted TV screen depicting a forest scene. Video game technology allowed participants to interact with and influence the forest elements by moving their hands and arms. Ten participants were invited to take part in one session lasting up to 15 minutes. The outcomes were measured before, during and after the intervention. Emotion and apathy were recorded using the **Observed Emotions Ratings Scale** (OERS)and **Person Environment Apathy Rating** respectively. Each participant’s engagement was coded into three types: (a) self-engagement: the resident engages in the activity without encouragement; (b) facilitated engagement: engagement in the activity is encouraged and supported by another person; and (c) no engagement: the resident is not engaging in the activity. The average duration of time spent in each type of engagement was used to describe engagement of the participant. The OERS ratings at pre, during and post-test were not compared, but the authors compared them to ratings previously established for people with dementia in an activity context. They reported scores were significantly higher for pleasure (p=0.008) and alertness (p<0.01) but not for sadness and anger. In addition, 50% of residents showed significantly higher fear. Apathy scores were significantly lower during the intervention (12.10), than before (18.30, *p*=0.01) or after (18.70, *p*=0.005). Participants spent an average of 10 minutes engaged in the Virtual Reality forest, and there were no differences in durations spent in the three different types of engagement. Conclusions: this pilot study suggests immersion in a Virtual Reality forest could be useful for increasing pleasure and alertness and reducing apathy in people with dementia, and a larger study should be undertaken to evaluate its effectiveness.

Powell *et al* (1979) used a **controlled crossover** design to determine whether taking part in an **indoor gardening programme** could increase **engagement** in 32 residents of a local authority nursing home in Hampshire, UK. Residents who habitually sat in one of two lounges were chosen to be involved. Gardening sessions were delivered on one day per week for 10 weeks (‘gardening days’). Residents could choose on each gardening day whether to attend the session or not. Whether or not they attended, their **engagement** level was observed at three minute intervals (total duration not stated). Participants were **recoded as engaged** if they were interacting with another person, using recreational materials, using materials connected with daily living activities or moving around using mobility equipment. If a participant did not attend the gardening session, engagement was measured in the lounges. Engagement was additionally measured for all participants on another day each week (‘non-gardening days’) whilst they were in the lounges. The number of participants who were engaged and the number being observed were calculated to produce a % engagement score for each day. The average attendance at the gardening sessions was 6.1 residents per session (range: 3-9). Over the course of the study, 12 different residents attended the gardening sessions. Engagement in the gardening sessions was consistently high, averaging 90%. Engagement in the lounges was consistently low, averaging 30% on non-gardening days and 31% on gardening days. A one-tailed t-test found that total engagement (including measures taken from participants in the lounge) was 43% on gardening days and 31% on non-gardening days (p<0.005). Engagement in the gardening sessions remained high (79%) at a 4-month post experimental timepoint. Conclusion: The engagement level of a selected portion of low-activity residents in a care home was significantly higher on days when an indoor gardening intervention was run.

Reynolds *et al* (2018) used a controlled crossover design to determine whether exposure to a **virtual nature experience** could reduce **agitation, anxiety** and **heart rate** for 14 people with dementia living in memory care units of an assisted living facility (location not specified). In one condition, residents watched a nature film depicting a mountain scene with a stream, accompanied by nature sounds. In the control condition, they viewed a generational movie. The film/movie were both shown on a 65 inch TV, in a small room, which was intended to increase feelings of presence in the experiences. Participants experienced both conditions three times in a counterbalanced design. There was a washout period of one day between conditions and one week between each of the three trials. Participants were required to stay in the room for at least ten minutes, but could stay longer if they wanted. The outcomes were measured with the **Observed Emotions Ratings Scale (OERS)**, **Agitated Behaviour Scale,** and **pulse oximetry** were administered at pre- and post-test. Repeated measures ANOVA was used to test for main effects and interactions. Heart rate significantly decreased from pre- to post-test in the nature condition (78.3 to 69.8 bpm, *p*=0.03) but not in the control condition (75 to 74.6 bpm, *p*=0.715). There was a significant time x treatment interaction effect (p=0.012). There were significant decreases in agitation and anger in both conditions, but no between-group differences. There was a non-significant decrease in pleasure in the nature condition (3.7 to 3.5, *p*=0.37), but a significant increase in the control condition (3.8 to 3.3, *p=*0.04), and a significant interaction between time and condition emerged (*p*=0.027). There were no significant effects for other OERS subscales of alertness, sadness or anxiety. Conclusion: virtual nature experiences may be a cost-effective way of increasing pleasure and decreasing stress for people with dementia.

Riddick (1985) used a **three group** **controlled pre/post quasi experiment** to determine whether owning a goldfish **aquarium** would significantly improve **blood pressure, happiness, anxiety, loneliness** and **leisure satisfaction** in a sample of 24 non-institutionalised elderly people. Residents of a low-income public subsidised apartment complex in the US were non-randomly assigned to aquarium, visits or control group. Aquarium participants had an aquarium containing two goldfish installed in their home. Researchers also visited the aquarium group ten times, once fortnightly for 25-40 minutes to help clean/maintain the tanks and talk about the fish. The visits group had visits only, and could choose what to do - most opted for socialising and watching television. The control group did not receive an aquarium or visits. **Memorial University of Newfoundland Scale of Happiness (MUNSH)**, the Trait A Scale of the **State Trait Anxiety Inventory (STAI), UCLA Loneliness Scale, Leisure Satisfaction Scale** and **blood pressure readings** were gathered pre and post-test. **The alpha level was set at 0.30**. T-tests were used for diastolic blood pressure, as baseline scores significantly differed. Diastolic bp significantly decreased from pre to post-test for the aquarium (82.00 vs 74.57, *t=*2.60, p=0.04) but not the visitor (72.25 vs 69.50, *t=*0.96, p=0.37) or control (78.86 vs 76.00, *t=*0.46, p=0.66). No between-groups comparisons were made. ANOVA was used for all other measures. Leisure satisfaction increased from pre to post test for the aquarium group relative to the other two groups (*F=*1.35, p=0.28). The aspect of the scale which improved most for aquarium vs the other groups was relaxation (*F=*3.17, p=0.06).The visitor group experienced a decrease in loneliness (*F=*2.86, p=0.08) compared to the other groups. There were no between group significant differences in any other measures, though systolic bp decreased in the aquarium and visitor groups, but not in controls. Conclusion: owning a fish aquarium may help reduce blood pressure and improve relaxation in some community-dwelling older adults. However, this study needs to be repeated with random allocation and testing at the 0.05 significance level.

Scott *et al* (2014) used a **three-group** **cluster-randomised ABA** design to determine the relative effects of a **biophilia installation** and a reminiscence installation on a range of health and wellbeing outcomes in 33 residents of three aged-care facilities in Australia. Each facility was randomly assigned to receive a **biophilia intervention,** a reminiscence intervention (objects, furniture and décor from the 1920s-1950s) or to control (no installation) for four weeks. **Satisfaction with living environment, satisfaction with opportunities for keeping occupied**, and **social engagement** were assessed using questionnaire items specifically constructed for this study. **Geriatric Depression Scale** short form (GDS-5), **Geriatric Anxiety Inventory** short form (GAI-SF), and the **Quality of Life - Alzheimer’s Disease** (QoL-AD) assessment tool were also employed. All tools were used at baseline (T1), during the intervention (T2) and during the two weeks after the intervention (T3) except QoL which was only measured at baseline. ANOVA, MANOVA and follow-up t-tests with Bonferroni correction were used to ascertain between group differences. A significant time x condition interaction occurred for social engagement (*F=*2.85, p<0.05). There were no differences at T1, but at T2, social engagement was significantly better in the biophilia (2.17) and reminiscence (2.42) groups than the controls (2.98) - NB lower scores mean better social engagement. Similar results were seen at T3. There were no significant differences between the biophilia and reminiscence conditions at any time point. There were no significant time or group effects for any other measure. However, there was a trend for satisfaction with living environment and satisfaction with opportunities for keeping occupied to increase during the intervention period in the biophilia group, but not the other groups. Conclusion: a 4-week biophilia or reminiscence installation significantly improved social engagement in a sample of residents living in aged-care facilities.

Talbot and Kaplan (1991) used a **cross-sectional questionnaire** to explore whether taking part in **indoor ‘compensatory’ nature activities** was associated with life satisfaction and residential satisfaction in a sample of 48 residents of an apartment complex for the elderly in Michigan, US. As part of a broader questionnaire about access to and perceptions of importance of nature, respondents were asked: “how much do you do each of the following?” 1) look at nature photographs or drawings, 2) watch nature programmes on TV, 3) use nature themes in decorating or when buying greetings cards. Answers were given on a 1-5 scale where 5=very often. They also answered the **Life Satisfaction Scale** and **Residential Satisfaction Scale**.Descriptive statistics were used. Mean scores were for the compensatory nature involvements were: 3.6/5 for looking at nature photographs, 3.5/5 for watching nature programs on TV, and 3.6/5 for using nature themes in decorating. T tests found that neither life satisfaction nor residential satisfaction scores were associated with the degree of compensatory nature involvements (data was not shown), but they were related to outdoor nearby nature. Conclusion: While one can read about nature, look at nature photographs or watch nature programs on TV, these compensatory activities do not relate to improved life or residential satisfaction.

Tse (2010) used a **cluster RCT** to explore whether an **indoor gardening programme** would impact on **activities of daily living (ADLs), socialisation**, **loneliness** and **life satisfaction** of 53 older people living in four nursing homes in Hong Kong. Two homes were assigned to the intervention group to take part in an indoor gardening programme once per week for 8 weeks. The other two homes were assigned to the control group, and received usual care including regular visits. **Life Satisfaction Index-A (Chinese version) (LSI-A), Revised UCLA Loneliness Scale (Chinese version), Lubben Social Network Scale (LSNS) and Modified Barthel Index (MBI)** were employed pre and post-test. Chi-square, Wilcoxon and spearman tests were variously used to test for differences between groups. The gardening group had significant pre-post improvements in all psychological measures: social network (19.27 vs 24.77, p<0.01), life satisfaction (11.73 vs 15.73, p<0.01), loneliness (41.38 vs 35.46, p<0.01), but not in ADL (107.19 vs 107.19, p=1.00). There were no significant pre/post-test changes in the control group. The post-test scores for the gardening group were significantly better than the control group for all measures (p<0.001) except ADL, for which there was no difference (p=0.06). Conclusions: Taking part in an 8-week indoor gardening programme significantly improved three psychological outcomes (but not activities of daily living) for older adults living in nursing homes.

Webster (2015) used a **one-group ABABB** design to study effects of **indoor plants** on **cognitive function** and **behaviour** of people with dementia in 11 residents of a continuing care facility in North Carolina, US. Plants were placed in frequently used common areas on day 1, were removed after day 6, replaced on day 8 and remained in place until final removal on day 15. Revised versions of the **Time and Change** and **Trail-making tests**, an attention task from the **Montreal Cognitive Assessment**, and a **colour recognition** task were used to assess cognitive function. **Dementia Care Mapping** was used to record behaviours, specifically 1) ill/wellbeing and 2) interactive (categorised into high potential, agitated and withdrawn). Measures were taken on day 1, 6, 8, 13 and 15. The overall **proportions** of improving scores and declining scores between various time points were compared. On the whole, cognition improved between day 1 and 6, and between day 8 and 15 (i.e. from baseline days to final plant days). Cognition was higher on day 6 (a final plant day) than day 8 (two days after plant removal). Ill/wellbeing and agitation remained relatively unchanged throughout. There was a tendency for high potential behaviours to decrease and withdrawn behaviours to increase during the plant periods. Conclusion: this study provides some support for the role of indoor plants to enhance cognitive ability in continuing care facility residents with dementia, but there seemed to be a negative association with behaviour.

Yao and Chen (2017) used a **controlled pre/post quasi-experiment** to test the effects of a **horticulture therapy programme** on **activities of** **daily living (ADLs), happiness, meaning of life and interpersonal intimacy** of 85 residents of 7 nursing homes in Kaohsiung, Southern Taiwan. The intervention group received 1 hour of horticulture therapy once per week for 8 weeks. The control group received usual care and was told they would have the opportunity to take part in the programme in due course (waiting list). **Barthel Index (BI), Chinese Happiness Inventory short form (CHI), Meaning of Life Scale (MLS)** and **Interpersonal Intimacy Scale (IIS)** were used pre and post-test. Paired t-tests were used to make within-group comparisons. ANCOVA was used to compare between-group post-test differences whilst controlling for pre-test differences as a covariate. The intervention group’s scores significantly improved from pre to post-test on ADL (62.44 to 67.32, t:-2.04, p=0.048), happiness (11.07 to 14.02, t:-5.09, p<0.001), meaning of life (74.49 vs 76.61, t:-3.25, p=0.002) and interpersonal intimacy (44.49 vs 48.39, t:-3.28, p=0.002). Meaning of life scores also improved for the controls (p=0.045) but decreased for happiness (p=0.002) and for ADL (p=0.059). Baseline scores between groups were not significantly different, but the intervention group had significantly better post-test scores than controls on ADL (70.80 vs 58.79, F=11.89, p=0.001), happiness (14.15 vs 7.86, F=59.18, p<0.001), interpersonal intimacy (49.24 vs 43.79, F=16.55, p<0.001) but not meaning of life (76.47 vs 76.19, F=0.11, p=0.738). Conclusion: Participating in horticultural activities significantly improved ADL, sense of happiness and interpersonal intimacy for older adults living in nursing homes.

Appendix 4. Detailed study outcomes

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| First author, year | Study design (quality appraisal score) | Intervention(s) | Control(s)/ comparator(s) | Tool | Within-group changes | Between-group differences |
| Dementia-specific outcomes | | | | | | |
| Cognition | | | | | | |
| D’Andrea, 2008 | RCT (strong) | Horticulture Therapy programme | Various regular scheduled activities e.g. music sessions, reminiscence discussions, socialising | TSIa  MDSa | TSI  I: 19.5 to 20.4  C: 21.5 to 19.5  MDS+  NR | I>C for pre to post score change (0.84 vs -2.0, t=5.7, p<0.0005) |
| Masuya, 2014 | CCT (weak) | Horticultural activity programme | Usual care | MMSEb | No main effect of time (F=1.41, p>0.05)  I: 23.1 to 23.4 (F=3.21, p>0.05)  C: 24.3 to 24.2 (F=0.13, p>0.05) | No main effect of condition (F=0.36, p>0.05)  No condition x time interaction (F=1.09, p>0.05) |
| Eggert, 2015 | Within-subjects quasi-experiment (weak) | Preferred nature images (photographs) | Preferred music | MoCAa | I: three of nine participants (33%) scored 2+ at pre-test compared with four of eight (50%) at post-test.  C: zero of four participants (0%) scored 2+ points at pre-test compared with two of six (33.4%) at post-test  No stats performed. | NR |
| Lee, 2008 | One-group pre/post (moderate) | Indoor gardening programme | N/A | HSD-Ra | 13.70 to 17.48 (*t=*12.044, p<0.001) | N/A |
| Webster, 2015 | One-group ABABB (weak) | Indoor plants | N/A | 1) Revised Time and Change Test  2) Revised Trail- making Test  3) MoCAa  4) Colour recognition task constructed for the study  Measures collapsed into one score | Reported as proportion of scores changing between time points.  From 1st baseline to 1st plant exposure: 28.8% of scores improved, 48.8% were unchanged, 22.5% declined.  Similar patterns seen between 2nd baseline and 2nd exposure and 2nd baseline and 3rd exposure.  No stats performed. | N/A |
| Agitation | | | | | | |
| Aslakson, 2010 | RCT (weak) | Nature videos | Music therapy | WAIb | Significant improvements in both groups (F=5.83, p<0.05)  I: effect size 0.46  C: effect size 0.33 | No significant time by condition interaction (F=0.04, p=0.84)  ES between I & C: ƞ2= 0.001 (no effect) |
| Reynolds 2018 | Quasi-crossover (moderate) | Nature film | Generational movie | Agitated behaviour scaleb | Significant improvements in both groups (both 15.7 to 14.9, p=0.003) | NS |
| Martin, 2011 | Cluster-randomised crossover (moderate) | Nature slide show looped corridor projection (180 photos) | Nursing home interiors slide show looped corridor projection (180 photos) | BARSb | No significant period effect (*t*=1.96, p=0.07) | No significant treatment effect (*t=*0.97, p=0.34)  A significant treatment-by-period interaction (*t*=2.52, p=0.02) was lost on adjustment for baseline measures |
| Cohen-Mansfield, 1998 | Quasi-crossover (weak) | Nature corridor enhancement | ‘Home’ corridor enhancement | CMAIb | NS reductions in frequency of agitation episodes occurred in three of four agitation categories for I, and two of four categories for C:  Physical, nonaggressive (n=5)  I: 27.76 to 17.60  C: 27.76 to 26.46  Physical, aggressive (n=2)  I: 4.23 to 2.50  C: 4.23 to 0.00  Verbal, nonaggressive (n=5)  I: 9.57 to 7.35  C: 9.57 to 9.74  Verbal, aggressive (n=6)  I: 16.26 to 16:00  C: 16.26 to 13.27 | NR |
| Lee, 2008 | One-group pre/post (moderate) | Indoor gardening programme | N/A | Modified CMAIb | 5.09 to 3.13 (*t* -4.002, p=0.001) | N/A |
| Chung, 2016 | One-group pre/post (weak) | Nature media presentations (DVDs) | N/A | Daily Nursing Records | No means reported but all changes were NS:  Hit a resident p=0.09  Hit a staff member p=0.58  Yelled/screamed p=0.34  Cursed/swore p=0.58  Experienced agitation to the point of interference with care  p=0.52  Threw self on the floor p=0.55 | N/A |
| Eggert, 2015 | Within-subjects quasi-experiment (weak) | Preferred nature images (photographs) | Preferred music | CMAIb | Only results from Q9 and Q10 (verbally aggressive behaviours) are discussed: results were mixed for both conditions | NR |
| Other dementia-related behaviours | | | | | | |
| Cohen-Mansfield, 1998 | Quasi-crossover (weak) | Nature corridor enhancement | ‘Home’ corridor enhancement | 1) Direct observation  2) Ambulatory tracking device.  Both reported as durationb | Direct observations:  Pacing/wandering  I: 137.60 to 137.11  C: 132.28 to 125.74  Trespassing  I: 8.61 to 7.32  C: 8.69 to 5.82  Exit-seeking  I: 0.34 to 0.20  C: 0.30 to 0.27  None were significant for either group  Ambulatory tracking  device: NS | NR |
| Aslakson 2010 | RCT (weak) | Nature videos | Music therapy | FBPa | Social Interactions  Means NR  I: effect size 0.48  C: effect size 0.57 | Social Interactions  F=0.93, p<0.05, direction NR  Task Performance and Problem Solving domains NS |
| Goto, 2014 | CCT (weak) | Indoor Japanese garden | Snoezelen room | 1) BACa  2) willingness to participate - recorded as either: absent, awake, asleep, leave, refuse.  3) frequency of verbalisationsa | NR | BAC: movement  I>C (X2=50.44, p<0.001)  Verbalisations  I>C (56% vs 24% of participants verbalised at least once)  Stayed awake  I>C (12.63 vs 2.5, t=11.18, p<0.001)  No stats performed for any other measure. |
| Webster 2015 | One-group ABABB (weak) | Indoor plants | N/A | Dementia Care Mapping | The proportion of participants exhibiting high potential behaviours tended to decrease and the proportion exhibiting withdrawn behaviours tended to increase when plants were in place. No stats performed. | N/A |
| Edwards, 2014 | One-group pre/post (moderate) | Fish aquarium | N/A | Nursing Home DBSb | Overall behaviour  67.2 to 58.2 (F=15.6, p<0.001)  Significant along Irrational, Uncooperative, Sleep and Inappropriate domains.  NS for Annoying and Dangerous domains. | N/A |
| Psychological wellbeing outcomes | | | | | | |
| Mood | | | | | | |
| Cohen-Mansfield, 1998 | Quasi-crossover (weak) | Nature corridor enhancement | Home corridor enhancement | LMBSa | Pleasure subscale  I: 1.05 to 1.11 (p<0.05)  C: means NR (NS) | NR |
| All other subscales NS for I & C |
| Reynolds 2018 | Quasi-crossover (moderate) | Nature film | Generational movie | OERS subscales:  Anxietyb  Angerb  Sadnessb  Pleasurea  Alertnessa | Anxiety: NS effects.  I: 1.9 to 1.5  C: 1.5 to 1.4  *p’s* not reported | NS interaction (*p*=0.268) |
| Anger: text states significant decreases of 1.4 to 1.1 (*p*=0.028) in both conditions, but text mismatches Table 1 data (possible typographical error). | NS |
| Sadness: NS effects. I: 1.3 to 1.3  C: 1.2 to 1.1  *p’s* not reported | NS interaction (p=0.263) |
| Pleasure: NS increase in I: 3.7 to 3.5 (p=0.37). Significant decrease in C: 3.8 to 3.3 (p=0.04) | Condition x time interaction effect p=0.027 |
| Alertness: NS effects. I: 4.9 to 4.9  C 4.9 to 4.9 | NS Interaction (p=0.765) |
| Moyle, 2018 | One group pre/post (weak) | Virtual Reality Forest | NA | OERS subscales:  Anxietyb  Angerb  Sadnessb  Pleasurea  Alertnessa | NR | Scores were significantly higher for pleasure (p=0.008) and alertness (p<0.01) but not for sadness and anger, when compared to scores previously established for people with dementia in an activity context.  50% of residents showed significantly higher fear |
| Affect | | | | | | |
| Barnicle & Midden, 2003 | CCT (moderate) | Horticulture activity programme | Usual care (waiting list) | Affect balance scalea | I: 5.42 to 7.61 (F=3.17, p=0.08)  C: 4.29 to 3.00 (F=0.70, p=0.40) | I>C for pre to post score change (F=6.78, p=0.01) |
| Perceived restoration | | | | | | |
| Kiyota, 2009 | Three-group cluster RCT (weak) | I(1) Indoor plants: active group  I(2) Indoor plants: passive group | No plants | Modified PRSa | No main effect of time (F=1.553, p=0.169) | No main effect of condition (F=1.094, p=0.359) |
| Condition x time interaction (F=2.115, p=0.023). No post hoc tests but appears due to late time point improvements in the active I(1) group |
| Happiness | | | | | | |
| Riddick, 1985 | Three-group CCT (weak) | Fish aquarium + researcher visits | C(1) Researcher visits only  C(2) Control (no aquarium, no researcher visits) | MUNSHa | I: -8.14 to -7.86  C(1): -2.00 to -4.38  C(2): -5.14 to -5.86  (F=0.94, p=0.41) | NR |
| Yao, 2017 | CCT (weak) | Horticulture Therapy programme | Usual care (waiting list) | CHIa | I: 11.07 to 14.02 (*t*= -5.09, p<0.001)  C: 9.84 to 7.98 (*t=* 3.27, p<0.002) | I>C for post-test scores (14.15 vs 7.86, F=59.18, p<0.001).  No difference in pre-test scores |
| Collins, 2008 | One-group pre/post (weak) | Horticulture activity programme | N/A | One item for SR happiness on a 1-4 Likert scalea | t1 to t2 *t*= 2.2, p=0.042  t1 to t3 *t*= -2.65, p=0.033  t2 to t3 *t=* -2.05, p=0.08 | N/A |
| Loneliness | | | | | | |
| Brown, 2004 | Cluster RCT (weak) | Indoor gardening programme | 20 minute visits from care staff | UCLA Loneliness Scale v3b | Group results NR but loneliness in all groups decreased over time (F=21.31, p<0.01) | NS |
| Tse, 2010 | Cluster RCT (moderate) | Indoor gardening programme | Usual care | Revised UCLA Loneliness Scale (Chinese version)b | I: 41.38 to 35.46 (p<0.001)  C: 42.56 to 42.44 (p=0.08) | I>C for post-test scores (35.46 vs 42.44, p<0.01).  No difference in pre-test scores |
| Riddick, 1985 | Three-group CCT (weak) | Fish aquarium + researcher visits | C(1) Researcher visits only  C(2) Control (no aquarium, no researcher visits) | UCLA Loneliness Scaleb | I: 36.14 to 40.00  C(1): 45.38 to 38.25  C(2): 42:00 to 42.29  (F=2.86, p=0.08) | NR |
| Mastery | | | | | | |
| Collins, 2008 | One-group pre/post (weak) | Horticulture activity programme | N/A | Pearlin & Schooler’s Mastery Scalea | t1 to t2 (*t*= -6.75, p=0.001)  t2 to t3 (*t=* -2.0, p=0.086  t1 to t3 (*t*= -4.07, p=0.005) | N/A |
| Apathy | | | | | | |
| Moyle, 2018 | One group pre/post (weak) | Virtual Reality Forest | NA | PEARb | t1 18.30  t2 12.10  t3 18.70  t1 to t2 (*p*=0.01)  t2 to t3 (*p*=0.005) | N/A |
| Anxiety | | | | | | |
| Scott, 2014 | Three-group cluster RCT (moderate) | Biophilia installation | C(1) Reminiscence installation  C(2) No installation | GAI-SFb | No main effect of time  I: t1 1.00, t2 1.20, t3 1.90  C(1): t1 2.57, t2 1.86, t3 2.00  C(2): t1 2.67, t2 3.16, t3 2.83  (all p>0.05) | No main effect of group (F<1.35, p>0.05) |
| Riddick, 1985 | Three-group CCT (weak) | Fish aquarium + researcher visits | C(1) Researcher visits only  C(2) Control (no aquarium, no researcher visits) | STAI (Trait A)b | Increase in all groups  I: 36.71 to 42.43  C(1): 42.13 to 47.75  C(2): 39.43 to 47.43  (F=0.09, p=0.92) | NR |
| Depression | | | | | | |
| Scott, 2014 | Three-group cluster RCT (moderate) | Biophilia installation | C(1) Reminiscence installation  C(2) No installation | GDS-5b | No main effect of time  I: t1 1.80, t2 2.00, t3 2.00  C(1): t1 2.38, t2 2.38, t3 2.75  C(2): t1 2.57, t2 2.43, t3 2.42  (all p>0.05) | No main effect of group (F<1.35, p>0.05) |
| Kiyota, 2009 | Three-group cluster RCT (weak) | I(1): Indoor plants: active group  I(2): Indoor plants: passive group | No plants | GDSb | No main effect of time (F=0.911, p=0.490) | No main effect of group (F=2.09, p=0.156)  No condition x time interaction (F=0.749, p=0.700) |
| Masuya 2014 | CCT (weak) | Horticulture activity programme | Usual care | GDS-15b | No main effect of time (F=2.92, p>0.05). However, significant improvement occurred in I but not C:  I: 5.7 to 3.6 (F=14.01, p<0.05)  C: 6.4 to 6.9 (F=0.31, p>0.05) | No main effect of condition (F=1.96, p>0.05) |
| Condition x time interaction (F=8.12, p<0.05) |
| Social outcomes | | | | | | |
| Social Engagement | | | | | | |
| Scott, 2014 | Three-group cluster RCT (moderate) | Biophilia installation | C(1) Reminiscence installation  C(2) No installation | Questionnaire item constructed for the study. Answered on a 1-5 Likert Scaleb | I: t1 2.85, t2 2.17, t3 2.72 (F=13.82, p<0.001)  C(1): t1 2.96, t2 2.42, t3 2.69, (F=3.62, p<0.05)  C(2): t1 2.95, t2 2.98, t3 3.20 (F=1.68, p>0.05) | Condition x time interaction (*F=*2.85, p<0.05).  Post-hoc tests:  I>C(2) at t2 (p<0.001) and t3 (p<0.01)  No difference in pre-test (t1) scores |
| I=C(1) at all time points |
| Brown, 2004 | Cluster RCT (weak) | Indoor gardening programme | 20 minute visits from care staff | Revised Social Provisions Scalea | Group means NR but socialisation improved in both groups over time:  Enhanced guidance (F=24.84, p<0.01)  Reassurance of worth (F=19.33, p<0.01)  Social integration F=28.15 p<0.01)  Enhanced reliable alliance (F=28.55, p<0.01) | NS |
| Tse, 2010 | Cluster RCT (moderate) | Indoor gardening programme | Usual care | LSNSa | I: 19.27 to 24.77 (p<0.01)  C: 18.26 to 18.33 (p=0.16) | I>C for post-test scores (24.77 vs 18.33, p<0.01).  No difference in pre-test scores |
| Interpersonal intimacy | | | | | | |
| Yao, 2017 | CCT (weak) | Horticulture Therapy programme | Usual care (waiting list) | IISa | I: 44.49 to 48.39 (*t*= -3.28, p=0.002)  C:46.30 to 44.57 (*t*= 1.60, p=0.116) | I>C for post-test scores (49.24 vs 43.79, F=16.55, p<0.001). No difference in pre-test scores. |
| Functional and physical outcomes | | | | | | |
| Sleep | | | | | | |
| Lee, 2008 | One-group pre/post (moderate) | Indoor gardening programme | N/A | 24h sleep diaries | Wake after sleep onset b   1. Duration: 75.22 to 54.65 mins (*t=* 2.781, p=0.011) 2. Frequency: 6.08 vs 2.21 occasions, *t=* 3.568, p= 0.002)   Nap ⱡ   1. Duration: 158.43 vs 85.87 mins (*t* 7.933, p= <0.001) 2. Frequency: 3.18 vs 1.95 occasions (*t*=6.480, p=<0.001)   Nocturnal Sleep Timea  440.48 vs 483.52 mins (*t=* -3.493, p=0.002)  Nocturnal sleep efficacya  85.09 vs 89.62% (*t=* -3.048 p=0.006) | N/A |
| Changes in sleep onset time, wake up time and total sleep time NS |
| Activities of Daily Living (ADL) | | | | | | |
| Masuya, 2014 | CCT (weak) | Horticultural activity programme | Usual care | ADL-20a | No main effect of time (F=1.43, p>0.05)  I: 30.0 to 30.8 (F=2.80, p>0.05)  C: 41.0 to 40.9 (F=0.13, p>0.05) | No main effect of condition (F=5.36, p>0.05)  No condition x time interaction (F=2.54, p>0.05) |
| Yao, 2017 | CCT (weak) | Horticulture Therapy programme | Usual care (waiting list) | BIa | I: 62.44 to 67.32 (*t*= -2.04, p=0.048)  C: 65.57 to 62.25 (*t*=1.94, p=0.059) | I>C for post-test scores (adjusted means 70.80 vs 58.79, F=11.89, p=0.001)  No difference in pre-test scores |
| Tse, 2010 | Cluster RCT (moderate) | Indoor gardening programme | Usual care | MBIa | I: 107.19 vs 107.19, (p=1.00)  C: 107.11 to 107.11 (p=1.00) | I=C for post-test scores (107.19 vs 107.11, p=0.06)  No difference in pre-test scores |
| Brown, 2004 | Cluster RCT (weak) | Indoor gardening programme | 20 minute visits from care staff | MDSa - Physical Functioning Scale | I: Means NR but significant improvements in Transferring (F=7.87, p<0.01), Eating (F=5.44, p=0.02), Toileting (F=6.28, p=0.01)  C: scores NR | NR |
| Nutrition | | | | | | |
| Edwards, 2002 | CCT (weak) | Fish aquarium | Scenic ocean picturec (waiting list) | Nutritional Intakea: weight of food consumed at mealtimes (grams) | I: increased 27.1% from pre to post-test. (*t*= -7.932, p<0.001)  C: NS change (*t*= –.882, *p* = .391). | NR |
| Edwards, 2013 | One-group pre/post (weak) | Fish aquarium | N/A | Food Intakea: weight of food consumed at mealtimes (grams) | Mean increase of 196.9g (25%, p<0.05)  Main effect of phase (F=85.7, p<0.001): the increase between pre-test and the end of the first 2 weeks was greater than between 2 weeks and the end of the study. | N/A |
| Body weight | | | | | | |
| Edwards, 2002 | CCT (weak) | Fish aquarium | Scenic ocean picturec (waiting list) | Weighta  (lbs) | I: mean increase of 1.65 Ibs (p<0.001)  C: NR | NR |
| Edwards, 2013 | One-group pre/post (weak) | Fish aquarium | N/A | Weighta  (lbs) | 158.4 to 160.6 (*t*=7.5, p=0.000 <0.05/3) | N/A |
| Physiological Outcomes | | | | | | |
| Reynolds 2018 | Quasi-crossover (moderate) | Nature film | Generational movie | Heart rateb (pulse oximetry) | I: 78.3 to 69.8 (p=0.03)  C: 75.0 to 74.6 (p=0.72) | Condition x time interaction (p=0.012) |
| DeSchriver, 1990 | RCT (weak) | Fish aquarium | C(1) Fish videosc  C(2) Videos of static | 1) Pulse rateb (brachial artery monitor)  2) Skin temperaturea (finger thermometer)  3) Muscle tension b (bicep EMG) | No significant within-group changes in any measure from pre to post-test observed for any group. | No significant differences in post-test scores between the three groups in any measure, controlled for pre-test scores. |
| Riddick, 1985 | Three-group CCT (weak) | Fish aquarium + researcher visits | C(1) Researcher visits only  C(2) Control (no aquarium, no researcher visits) | Blood pressureb (Sphygmomanometry) | Diastolic BP  I: 82.00 to 74.57 (*t=*2.60, p=0.04)  C1: 72.25 to 69.50 (*t=*0.96, p=0.37)  C2: 78.86 to 76.00 (*t=*0.46, p=0.66) | NR |
| Systolic BP  NS changes for any group (F=0.31, p=0.74) |
| Goto, 2014 | CCT (weak) | Indoor Japanese garden | C(1) Snoezelen room  C(2) Participant’s bedroom | Pulse rateb (finger plethysmography) | I: consistent decline of 0.15-0.2 bpm throughout sessions  C1: variable but average increase of 0.06 bpm throughout sessions.  No stats performed. | I>C2 for the final 6 min of recording (p=0.034)  C1=C2 for the final 6 min of recording (p=0.34).  No direct I to C1 comparison |
|  | General health, wellbeing and satisfaction outcomes | | | | | |
| Engagement | | | | | | |
| Powell, 1979 | CCT (weak) | Gardening days (optional indoor gardening session) | Non-gardening days | Constructed for study: is participant engaged? yes/no Converted to % engaged as a proportion of total sample. | I: engagement ranged from 34-51% across sessions  C: engagement ranged from 20-41% across sessions  No within-group stats performed but levels varied for both conditions | I>C engagement level averaged 43% on gardening days and 31% on non-gardening days (p<0.005) |
| Aslakson, 2010 | RCT (weak) | Nature videos | Music therapy | Tool constructed for studya | I: 3.4 to 2.3, effect size 0.59 (moderate).  C: 3.3 to 4.5, effect size 0.94 (high) | I<C for pre to post-score change (F=24.54, p<0.01)  Effect size between I & C: ƞ2=0.392 |
| Eggert, 2015 | Within-subjects quasi-experiment (weak) | Preferred nature images (photographs) | Preferred music | IDEASa | I & C: slight score improvements between 0.7 and 1.8 from pre to post-session every week. No stats performed. | NR |
| Moyle, 2018 | One group pre/post (weak) | Virtual Reality Forest | NA | Duration spent engaged in the activitya coded as:   * No engagement: the participant is not engaging in the activity * Self-engagement: participant is engaging without encouragement * Facilitated engagement: participant is engaging with encouragement from staff | Total time engaged with the Virtual Reality Forest ranged from 8.03 to 12.30 min, with an average of 10.22 min (*SD* = 1.07). There were no significant differences in durations spent in each type of engagement:  Not engaged: 4.45 min (*SD* = 2.4)  Facilitated: 3.33 min (*SD* = 1.57)  Self-engaged: 2.44 min (*SD* = 2.11) | NA |
| Perceived health | | | | | | |
| Collins, 2008 | One-group pre/post (weak) | Horticulture activity programme | N/A | SR healtha | t1 to t2 (*t*= -4.12, p=0.001)  t2 to t3 (*t=* -0.75, p=0.02)  t1 to t3 (*t*= -3.99, p=0.005) | N/A |
| Perceived wellbeing | | | | | | |
| Kieffer, 2014 | Cross-sectional (weak) | Representational elements of nature (photographs) | N/A | Questionnaire constructed for the study containing five items relating to perceived wellbeinga   1. Refreshed-exhausted 2. Distracted – attentive 3. Relaxed – harried 4. Irritable – patient 5. Comfortable – uneasy | Images containing water perceived as better for wellbeing than images without water on all 5 scales.  Images containing natural materials perceived as worse for wellbeing than pictures without natural materials on 4/5 scales.  Results for images containing fire and botanical motifs were mixed.  No stats. | N/A |
| Quality of Life | | | | | | |
| Masuya, 2014 | CCT (weak) | Horticultural activity programme | Usual care | VASa | Satisfaction with life  I: 75.6 to 91.1 (F=28.00, p<0.05)  C: 78.8 to 81.1 (F=2.29, p>0.05) | Satisfaction with life  Condition x time interaction (F=16.46, p<0.05) |
| No significant changes for either group in any of the other 6 subscales | NS differences in the other 6 subscales |
| Meaning of Life | | | | | | |
| Yao, 2017 | CCT (weak) | Horticulture Therapy programme | Usual care (waiting list) | MLSa | I: 74.49 to 76.61 (t= -3.25, p=0.002)  C: 73.89 to 76.07 (t= -2.07, p=0.045) | I=C for adjusted post-test scores (76.47 vs 76.19, F=0.11, p=0.738)  No difference in pre-test sores |
| Life satisfaction | | | | | | |
| Tse, 2010 | Cluster RCT (moderate) | Indoor gardening programme | Usual care | LSI-Aa (Chinese version) | I: 11.73 to 15.73 (p<0.01)  C: 11.56 to 11.67 (p=0.08) | I>C for post-test scores (15.73 vs 11.67, p<0.01). No difference in pre-test scores |
| Satisfaction with opportunities to keep occupied | | | | | | |
| Scott, 2014 | Three-group cluster RCT (moderate) | Biophilia installation | 1. Reminiscence installation 2. No installation | Questionnaire constructed for the study: answers on 1-5 Likert Scalesa | No main effect of time  I: t1 2.10, t2 1.90, t3 2.20  C(1): t1 2.38, t2 2.38, t3 2.75  C(2): t1 2.57, t2 2.43, t3 2.42  (all p>0.05) | No main effect of condition (F<1.35, p>0.05) |
| Leisure satisfaction | | | | | | |
| Riddick, 1985 | Three-group CCT (weak) | Fish aquarium + researcher visits | 1) Researcher visits only  2) Control (no aquarium, no researcher visits) | LSSa | I: 63.57 to 74.00  C(1): 68.50 to 74.50  C(2): 67.00 to 65.29  (F*=*1.35, p=0.28) | NR |
| Satisfaction with living environment | | | | | | |
| Scott, 2014 | Three-group cluster RCT (moderate) | Biophilia installation | 1. Reminiscence installation 2. No installation | Questionnaire constructed for the study: answers on 1-5 Likert Scalesa | No main effect of time  I: t1 2.30, t2 1.70, t3 2.20  C(1): t1 2.13, t2 2.13, t3 2.13  C(2): t1 2.14, t2 2.14, t3 2.29  (all p>0.05) | No main effect of condition (F<1.35, p>0.05) |

Appendix 4. Detailed study outcomes. Scores reported as group means unless otherwise stated. C control/comparator; I intervention; N/A not applicable; NR not reported; NS not significant (no means provided); t1 baseline/pre-test scores; t2 mid-intervention or post-test scores; t3 post-test or follow-up scores; aincreasing scores = improvement in outcome; bdecreasing scores = improvement in outcome; ccondition was a control group in the individual study but is considered as an intervention in the current review.Colour coding: in the within-groups column, green indicates the intervention group (but not the control/comparator group, if applicable) was significantly effective over time; yellow indicates the intervention group had no significant effect over time; red indicates the control/comparator group (but not the intervention group) was significantly effective over time. In the between-groups column, green indicates a significant difference favouring intervention; yellow indicates no between-group differences; red indicates a significant difference favouring control/comparator. In either column, no colour indicates results were inconsistent/mixed, or direction of effect(s) were unclear, or some results were not reported, or the study only reported descriptive statistics, or the test was not performed or not applicable.

ADL Activities of Daily Living; BAC Behavioural Assessment Checklist; BARS Brief Agitation Ratings Scale**;** BI Barthel Index; CCT controlled clinical trial; CHI Chinese Happiness Inventory; CMAI Cohen-Mansfield Agitation Inventory; DBS Disruptive Behaviour Scale; FBP Functional Behaviour Profile; GAI-SF Geriatric Anxiety Inventory short form; GDS Geriatric Depression Scale; GDS-5 Geriatric Depression Scale short form; HDRS-R Hasegawa Dementia Rating Scale – Revised; IDEAS Individualized Dementia Engagement and Activities Scale; IIS Interpersonal Intimacy Scale; LMBS Lawton’s Modified Behaviour Stream; LSI-A Life Satisfaction Index-A; LSNS Lubben Social Network Scale; LSS Leisure Satisfaction Scale; MoCA Montreal Cognitive Assessment; MBI Modified Barthel Index; MDS Minimum Data Set; MDS+ Minimum Data Set Plus; MLS Meaning of Life Scale; MMSE Mini Mental State Exam; MUNSH Memorial University of Newfoundland Scale of Happiness; N/A not applicable; OERS Observed Emotions Ratings Scale; PEAR Person Environment Apathy Rating; PRS Perceived Restoration Scale; QoL Quality of Life; RCT randomised controlled trial; RSS Residential Satisfaction Scale; SR self-reported; STAI State-Trait Anxiety Inventory; TSI Test for Severe Impairment; VAS Visual Analogue Scale; WAI Wisconsin Agitation Inventory.