**Full Methods:** Extended methods section, including more detailed search, screening, and quality assessment strategies, as well as levels of agreement for screening and risk of bias rating for transparency.

The systematic literature review protocol was pre-registered on Prospero (ID: CRD42022347606). We adhered to the Non-Interventional, Reproducible, and Open Systematic Review (NIRO-SR) pre-registration and reporting guidelines (V1)26 to encourage transparency and reduce bias. Whilst papers that explored eye movements during naturalistic locomotion and balance tasks were collected as part of the search strategy described in the pre-registration and below, these papers were excluded from this review and will instead be summarised in a separate review. There were no other deviations from the protocol.

**Search Protocol**

The systematic literature search was conducted on 25th August 2023. To account for the multidisciplinary nature of the research question and ensure the completeness of the review, records were identified through: PsycInfo (via Ovid), Medline (via Ovid), Scopus, and Web of Science. Databases were accessed via the University of Oxford. The search strings applied were adapted from the search strings used in our earlier review exploring naturalistic eye movements in Alzheimer’s disease, which were created with the aid of a librarian.23 The search strings encompassed search terms relating to both the populations (PD, healthy older adults) and tasks (naturalistic eye movement tasks) of interest. As the databases utilised did not support the use of dictionary terms, only free-text search terms were used and remained consistent across each database (see Supplementary Material 1 for the full search strategy).

To account for relevant records that may have been missed from the database searches, we completed forward and backward citation tracking of included papers using Google Scholar. Additionally, as grey literature is often excluded from large bibliographic databases, we searched for relevant pre-prints using PsyArXiv, MedArXiv, and bioArXiv, to prevent publication bias. The titles of all pre-prints were searched on 23rd October 2023 with combinations of the following search terms: “eye movements” and “Parkinson's”. Full texts were screened if a record appeared as though it could be relevant from its title and abstract.

The .ris files downloaded during the final search of each database were exported into CADIMA, an online open-access tool designed to facilitate each stage of conducting a systematic review.27,28 The records were de-duplicated using CADIMA’s automated duplicate removal tool and those records identified as potential duplicates were checked by hand. Of the 2,309 records identified from the final database searches, 1,030 (44.61%) duplicates were identified and removed prior to screening.

**Inclusion and Exclusion Criteria**

Screening was conducted in two phases: (1) title and abstract screening, and (2) full-text screening. The screening criteria applied during the title and abstract screening were: (1) peer-reviewed original studies reported in the English language, (2) the inclusion of an idiopathic PD group without psychiatric or neurological comorbidities, (3) the use of a naturalistic eye movement task, except tasks conducted during locomotion or assessing balance (see Supplementary Material 2 screening instructions). Full text screening criteria were more comprehensive and were: (1) peer-reviewed original studies reported in the English language, (2) the inclusion of an idiopathic PD group without psychiatric or neurological comorbidities and an appropriate HC group, (3) the use of a naturalistic eye movement task, (4) reported statistics for the comparison of eye movements between PD and HC (see Supplementary Material 2).

Psychiatric comorbidities, such as depression and anxiety, are common in PD,32,33 along with additional neurological conditions, such as rapid eye movement sleep behaviour disorder34 that can lead to disruptions in eye movements.35 As the aim of this review is to explore naturalistic eye movements in PD, we excluded PD patients with reported psychiatric and neurological comorbidities, as comorbidities may independently affect eye movements. Moreover, we excluded studies that only included PD patients with reported comorbid mild cognitive impairment (PD-MCI) or Parkinson’s disease dementia (PDD), as we are interested in whether naturalistic eye movement tasks can be used to identify PD in the early stages of the disease before any significant cognitive impairment occurs. Finally, research has observed significant differences in eye movements between people with idiopathic PD and other aetiologies of parkinsonism.29-31 Therefore, papers that only recruited patients with vascular parkinsonism, drug-induced parkinsonism, multiple systems atrophy (MSA), progressive supranuclear palsy (PSP), or corticobasal degeneration were excluded from the review. Whilst studies that recruited both idiopathic PD patients and atypical parkinsonian patients were included, only the idiopathic PD group will be discussed in this review.

The definition of naturalistic eye movement tasks adhered to was consistent with that used in an earlier review exploring naturalistic eye movements in Alzheimer’s disease.23 Naturalistic eye movement tasks were defined as tasks that either (a) incorporate naturalistic stimuli (e.g., emotion recognition task, goal-directed visual search using a naturalistic scene), (b) allow free unrestricted visual exploration of stimuli that are present for a minimal duration of 5s (e.g., visual paired comparison task, free viewing task), or (c) tasks identical to (e.g., reading, watching video clips) or closely mirror (e.g., virtual reality) tasks undertaken in a normal daily life setting. We regard prosaccade tasks involving naturalistic stimuli, such as a photograph, as a naturalistic paradigm, as they replicate eye movements regularly executed in daily life. In contrast, as antisaccade tasks do not require fixation on an object of interest, and instead requiring inhibition of eye movements away from an object of interest, we argue that these eye movements are counterintuitive to those performed in daily life and antisaccade tasks were therefore excluded

**Screening**

Two raters (MG, JH) independently screened all records during both phases of screening. Inter-rater reliability of title and abstract screening was assessed for a randomly selected sample of 26 (2%) studies and was deemed to be good (κ=0.65) before any further records were assessed. Similarly, inter-rater reliability in the full text screening process was excellent (κ=0.81) in a randomly selected sample of 18 (10%) full-text papers. The level of agreement between the two raters was 79% following title and abstract screening and 88% following full-text screening. Only 24% of title and abstract screening inconsistencies and 23% of full-text screening inconsistencies affected whether the record would have been included in the review (see Supplementary Materials for decision log). Of the 1,279 records screened, 1256 (98%) were excluded due to not meeting our inclusion criteria (see Fig. 2 for pictorial depiction of the numbers of records excluded at each stage of the review). Overall, 30 papers are included in the review with 23 being identified during full text screening and a further 7 being added during citation (N=6) and grey literature (N=1) searches. All records meeting the full text screening criteria were checked for retraction using the Retraction Watch Database on 23rd October 2023 (<http://retractiondatabase.org/>) with none of the included records being found within this database.

**Data Extraction**

Data extraction for each record was conducted by a single reviewer (MR) using Excel (see Supplementary Materials for the data extracted from each included paper). All extracted data were checked by another reviewer (MP) to ensure no missing data and that the extracted data were correct.

**Quality Assessment**

A quality assessment of each study that passed through full-text screening was conducted to inform the interpretation of individual studies. As no appropriate tool currently exists for assessing the quality of non-interventional research, the Downs and Black risk of bias checklist37 was adapted to be more appropriate for assessing the quality of the non-interventional research included in our review (see Supplementary Material 3). This tool has been commonly modified for assessing the quality of non-interventional research in previous systematic reviews (e.g., Readman et al.23 and Luo et al.38). The ratings for each criterion were updated from ‘yes’ and ‘no’ (scoring 1 and 0 points, respectively) to ‘yes’, ‘partially’, and ‘no’ (scoring 2, 1, and 0 points, respectively). Additionally, we changed the terminology to be more appropriate to non-interventional research (for example, ‘patients’ to ‘participants’, ‘lost to follow-up' to ‘lost to drop-out or exclusion’). Items 8, 12, 13, 14, 15, 17, 19, 21, 22, and 24 were removed as these could not be edited to apply to non-interventional research.

Each record was independently assessed by two reviewers (MG, JH) with any inconsistencies in ratings being resolved by a third reviewer (MP). The level of agreement between the two raters was 80.98% (see Supplementary Materials for a log of decision making). The maximum score for our modified risk of bias checklist was 34, whilst the mean score of all studies was 25.0 (range=16-31).