Morph-dependent fitness and directional change of morph frequencies over time in a Dutch population of Common buzzards Buteo buteo

Kappers, Elena Frederika; de Vries, Christiaan; Alberda, Anneke; Kuhn, Sylvia; Valcu, Mihai; Kempenaers, Bart; Both, Christiaan

Published in:
Journal of Evolutionary Biology

DOI:
10.1111/jeb.13675

Publication date:
2020

Received: 27 March 2020 | Revised: 30 June 2020 | Accepted: 30 June 2020

DOI: 10.1111/jeb.13675

RESEARCH PAPER

Journal of Evolutionary Biology eseb WILEY

Morph-dependent fitness and directional change of morph frequencies over time in a Dutch population of Common buzzards *Buteo buteo*

Elena Frederika Kappers[1,2] | Christiaan de Vries[3] | Anneke Alberda[3] | Sylvia Kuhn[2] | Mihai Valcu[2] | Bart Kempenaers[2] | Christiaan Both[1]

[1]Conservation Ecology Group, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands

[2]Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany

[3]Noormanstrjitte, Wijnjewoude, The Netherlands

Correspondence
Bart Kempenaers, Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany.
Email: b.kempenaers@orn.mpg.de

Funding information
Max-Planck-Gesellschaft
Projekt DEAL, Open access funding enabled and organized by Projekt DEAL.

Abstract

How genetic polymorphisms are maintained in a population is a key question in evolutionary ecology. Previous work on a plumage colour polymorphism in the common buzzard *Buteo buteo* suggested heterozygote advantage as the mechanism maintaining the co-existence of three morphs (light, intermediate and dark). We took advantage of 20 years of life-history data collected in a Dutch population to replicate earlier studies on the relationship between colour morph and fitness in this species. We examined differences between morphs in adult apparent survival, breeding success, annual number of fledglings produced and cumulative reproductive success. We found that cumulative reproductive success differed among morphs, with the intermediate morph having highest fitness. We also found assortative mating for colour morph, whereby assortative pairs were more likely to produce offspring and had longer-lasting pair bonds than disassortative pairs. Over the 20-year study period, the proportion of individuals with an intermediate morph increased. This apparent evolutionary change did not just arise from selection on individual phenotypes, but also from fitness benefits of assortative mating. The increased frequency of intermediates might also be due to immigration or drift. We hypothesize that genetic variation is maintained through spatial variation in selection pressures. Further studies should investigate morph-dependent dispersal behaviour and habitat choice.

KEYWORDS

assortative mating, *Buteo*, colour morph, colour polymorphism, plumage colour, reproductive success, survival

1 | INTRODUCTION

Identifying the processes that maintain genetic variation in populations over time is fundamental to evolutionary biology, as evolutionary responses are often based on the standing genetic variation (Lewontin, 1974). One wide-spread type of genetic variation in animals is colour polymorphism: within-population variation in appearance across individuals independent of age and sex

(Huxley, 1955). Understanding the persistence of genetically determined phenotypic polymorphisms requires measuring the covariation between the genetic part of the trait and fitness (Roulin, 2004).

Several mechanisms have been proposed to explain the maintenance of polymorphisms, and one of these is balancing selection. Negative frequency-dependent selection is a form of balancing selection where the rare allele has higher fitness than the more common alleles (Sinervo & Calsbeek, 2006; Smith, 1982). Another rarely observed form of balancing selection is overdominance, where heterozygotes have higher fitness than both homozygotes (Allison, 1964; Knief et al., 2017).

In birds, 3.5% of all species show plumage colour polymorphisms, which are often genetically determined (Galeotti, Rubolini, Dunn, & Fasola, 2003). Colour polymorphisms are particularly common in raptors (30% of species; Fowlie & Krüger, 2003; Hugall & Stuart-Fox, 2012). Among raptors, the common buzzard, *Buteo buteo*, is one of the most variable species in terms of plumage colour. Individuals vary along a dark–light continuum, but for practical reasons this variation has often been categorized into three morphs: dark, intermediate and light (Kappers et al., 2017). Using parent–offspring comparisons, we previously showed that plumage colour is highly heritable (82%), independent of sex and not influenced by environmental factors such as shared nest environment and parental identity (Kappers et al., 2018). Because the variation in these morphs has such a strong genetic basis, the covariation between phenotype and fitness can be considered as direct selection on the genetic component of the polymorphism. Previous studies on a German population provided evidence that the polymorphism could be explained by a one-locus two-allele inheritance system, with the intermediate morph being the heterozygote (Krüger, Lindström, & Amos, 2001). Krüger et al. showed that both lifetime reproductive success and adult survival were higher for the intermediate morph than for the light and dark morphs (see also Jonker, Chakarov, & Krüger, 2014). The authors concluded that heterozygote advantage maintains the colour polymorphism in this species (Krüger et al., 2001).

Fitness differences among morphs were also investigated in the related Swainson's hawk, *Buteo swainsoni* (Briggs, Collopy, & Woodbridge, 2011). Similar to the common buzzard, this species also shows continuous colour variation, which has been categorized in three morphs. Using 32 years of breeding data, Briggs et al. (2011) found no evidence that intermediate individuals (presumed heterozygotes) had higher fitness; there were no differences in any of the examined fitness components between the morphs. Therefore, Briggs et al. (2011) excluded both frequency-dependent selection and heterozygote advantage as mechanisms maintaining the colour polymorphism in this species.

Our aim is to replicate the study on fitness consequences associated with plumage colour morphs in the common buzzard. Replications of this type are relatively rare (Nakagawa & Parker, 2015), especially when carried out in the wild and on long-lived species, because they require long-term data sets. Replication is essential to validate findings, and it is a basic requirement for the advancement of any field of research to be able to generalize. Because outcomes in ecological and evolutionary studies often rely on specific ecological settings, it is likely that such replications will yield different outcomes, thereby showing the importance of studying the ecological causes underlying selection. An additional reason to replicate the previous study is that we recently showed that the colour polymorphism of the buzzard does not fit the originally proposed one-locus two-allele system of inheritance, but rather should be considered as a polygenic quantitative trait with high heritability (Kappers et al., 2018).

Here, we use a 20-year study of a population of common buzzards from The Netherlands to replicate the empirical findings on fitness consequences of plumage morph from the original publications (Jonker et al., 2014; Krüger et al., 2001). Specifically, we investigate differences among morphs in adult survival, annual reproductive rates and cumulative reproductive success.

An intriguing conclusion from the original study was that buzzard mate choice was maladaptive (Krüger et al., 2001), because pairs showed assortative rather than disassortative mating. This was supposedly maladaptive because to produce offspring with the highest fitness (the intermediate morph), light or dark individuals should mate with the opposite morph. Therefore, we also describe the mating patterns in relation to morph in our population and the fitness consequences of different mating combinations.

2 | MATERIALS AND METHODS

2.1 | Study site and population

We studied common buzzards from 1996 onwards in Friesland, The Netherlands (53°04′N, 6°13′E). The study site encompasses a 5,724-ha area with 1,400 ha of forested patches. The larger patches are spruce, pine and larch-dominated (~1,000 ha), whereas the smaller patches (~ 400 ha) are oak-dominated. The study area contained on average 81 ± 14 *SD* breeding pairs/year (range: 57–110). In each year, all territories were visited in late winter to determine whether they were occupied by a breeding pair. Breeding performance of each pair was assessed by multiple observations, including observations from the ground (nest building) and nest checks. Here, we use data from a 20-year period (1996–2015), during which all breeding buzzards were colour-scored for overall plumage, using a seven-morph scale ranging from very dark to very light (Kappers et al., 2017).

Individuals were identified based on plumage colour and pigmentation patterns scored from direct sightings in the field, photographs, captures (N = 90), and—in most cases—from collected moulting feathers, combined with the location of the observation (266 adult females and 244 adult males). Each year, we tried to collect moulting feathers of all females during incubation around the nest and for all males in their territory after the breeding season. Individual identification was based on visual comparison of the highly diverse colour patterns with collected feathers from previous years. We confirmed this individual assignment through genetic profiling with microsatellites using DNA from the shafts of a subset of collected moulted

feathers. Identification based on feather phenotype was correct for 99% of 199 analysed feathers (see Appendix S1).

For analyses, we grouped individuals into dark, intermediate and light morphs following a three-morph scheme (see Kappers et al., 2017). This allowed a direct comparison with previous studies in Germany (Krüger et al., 2001).

As our interest is also in potential ecological drivers of selection on colour morphs, we considered two environmental covariates that may affect annual variation in fitness components: (a) The North Atlantic Oscillation (NAO) index for the months December through March, which is indicative of the severity of the winter; positive values are typically associated with wetter and milder weather over western Europe, whereas negative values indicate drier and colder weather (updated from Jones, Jonsson, & Wheeler, 1997), and (b) an annual vole index, determined by the sum of the number of common vole *Microtus arvalis* holes in western Drenthe (approximately 20 km south from our study site) that were re-opened 24 hr after closing them in 35 grassland plots of 1 × 1 m in March and August (Bijlsma, 2016). Common voles vary strongly in abundance between years and are the primary food source for common buzzards (Dare, 2015). The NAO and vole index were only weakly correlated (Figure S1).

2.2 | Adult apparent survival

We used Cormack–Jolly–Seber (CJS) models to analyse whether survival of breeding adults observed between 1996 and 2015 was associated with morph. The data set included 266 individual females and 244 individual males (155 dark, 253 intermediate and 102 light individuals). The CJS models separate the survival probability from the resighting probability using a maximum likelihood approach. We analysed the sexes separately because high mate fidelity in buzzards increases the probability of observing a pair, such that male and female partners are nonindependent observations. We constructed our models using the program MARK (White & Burnham, 1999) with package RMARK (Laake, 2013) in R (R Core Team, 2016).

Our initial model for each sex included morph and year. First, we assessed the fit of these general models by performing goodness-of-fit (GOF) tests using the program RELEASE (Burnham, Anderson, White, Brownie, & Pollock, 1987). The GOF of the CJS models (test 2 and 3) was satisfactory (males: $\chi^2_{112} = 198.29$, $p < .0001$; females: $\chi^2_{112} = 165.80$, $p = .0007$). We found no indication of significant overdispersion (GOF test: $\hat{C}_{males} = 1.77$, $\hat{C}_{females} = 1.48$), but we corrected for the lack of fit of the model to the data by adjusting $\hat{C}$ from 1.0 to 1.77 for males and to 1.48 for females.

We estimated the parameters apparent survival (φ) and encounter probability (p). We used a hierarchical modelling approach, retaining only the best-ranked models from the previous step—based on Aikaike's Information Criterion (AIC)—before considering a new suite of covariates (Burnham & Anderson, 2002). As p was considered a nuisance parameter, we modelled it first to obtain the best fit for resighting probability. We added the factors morph (m) and year (t) to account for potential differences in detectability between the three morphs as well as among the years. We compared models based on a combination of ΔAICc and model complexity (number of parameters) following Burnham and Anderson (2002). Models that only added complexity to a simpler model but did not improve the fit (usually falling within two AICc values) were not considered competitive (Arnold, 2010). As there was no support for a difference between morphs in resighting probability, we only kept the model with year in subsequent analyses (see Table S1).

To model survival (φ), we used a first set of models with morph and year and their interactions ($m \times t$) as factors. We fitted all five possible models to the data. Subsequently, we added environmental variables that might affect survival probability. The new continuous variables included (a) the NAO-index, (b) the vole index and (c) the average number of fledged chicks from the previous year as a measure of how stressful the breeding season had been. We also added several biologically plausible interactions between these variables (see Table S2 for full set of 24 models). Figure S2 shows yearly variation in the ecological variables for the period 1996–2015. We did not include minimal age (or breeding career length), because age together with year severely reduces the degrees of freedom.

2.3 | Morph-assortative mating

We assessed the level of morph-assortative mating by calculating the Pearson's correlation between the colour morphs of pair members (for this we encoded 1 as dark, 2 as intermediate and 3 as light). We did this for two data sets. First, we considered all unique pair combinations of known colour morph from the entire study period ($N = 400$). Second, we considered all breeding pairs for each year of the study (including repeats; $N = 1{,}566$). For all breeding pairs, we assessed the level of morph-assortative mating across and within years. Most of the breeding attempts in multiple years were with the same partner (females: 68%, males: 63%). In the remaining cases, individuals had multiple mates during their stay in the population (females: 19%, 10%, 2% and 1% with 2, 3, 4 and 5 mates, respectively; males: 22%, 9%, 4%, 1% and 1% with 2–6 mates, respectively). The significance of the Pearson's correlation coefficients was tested using a resampling procedure. The values of the morph of all individuals were randomized $N = 100{,}000$ times (over the years when comparing unique pair combinations with all breeding pairs; within year when considering breeding pairs repeatedly) and for each randomization we calculated the correlation coefficient. The significance level of the actual correlation coefficient is then given by (n*2)/N, where n is the number of randomized values that are equal to or more extreme than the observed correlation.

We tested whether pairs with different degrees of assortment with respect to plumage morph differed in the duration of their pair bond. For all unique pair combinations of known colour morph ($N = 400$), the level of assortment by morph was defined in three categories: '2' for pairs where both members have the same morph (dark–dark, intermediate–intermediate, light–light), '1' for intermediate-dark or

intermediate-light pairs and '0' for dark-light pairs. We used a linear mixed model with pair duration in years (log_{10}-transformed) as the response variable and with the degree of assortment (factor with 3 levels) as the independent variable. We included as random intercepts the first year of breeding ($N = 20$, which also accounts for shorter pair durations in more recent years), female identity ($N = 264$) and male identity ($N = 242$). Results were back-transformed for illustrative purposes.

2.4 | Measures of reproductive success

For each breeding season, we examined reproductive success of all territories where the productivity was known and the morph of both adults had been scored ($N = 1,359$). We defined yearly reproductive success as the number of fledglings produced in that year, which varied between zero and four (mean ± *SD*: 0.9 ± 1.0, including all territories; 1.8 ± 0.8, $N = 732$, excluding unsuccessful nests and non-breeding pairs). Nestlings were considered fledged if their presence was recorded in the natal territory after the expected fledging date. For pairs that had two nesting attempts in the same breeding season ($N = 125$ pairs, a second attempt only occurred after failure of the first attempt), we only considered the last nest, because the first attempts were unsuccessful.

We modelled variation in yearly reproductive success with a GLMM using package *lme4* (Bates, Mächler, Bolker, & Walker, 2015) in R (version 3.3; R Core Team, 2016) using a Poisson distribution, a log-link and a Laplace approximation. As explanatory variables, we added morph of both attendant adults, the degree of assortment by morph of the pair (factor with three levels, see above) and 'disturbance' (factor with three levels: nest disturbed by humans, $N = 231$; nest take-over attempt by Egyptian goose *Alopochen aegyptiaca*, $N = 17$; no evidence of disturbance, $N = 1,111$). We also added female identity ($N = 259$), male identity ($N = 239$) and year ($N = 20$) as random intercepts.

Additionally, we calculated cumulative reproductive success for both males ($N = 244$) and females ($N = 266$), as the total number of fledglings produced during the adults' presence in the population. This variable ranged from 0 to 38 for both sexes (mean ± *SD* = 5.2 ± 6.4, for males; 4.7 ± 6.4 for females). For 60% of all individuals in the analysis, cumulative reproductive success (CRS) equals lifetime reproductive success (LRS)—assuming that individuals that were not observed during three consecutive breeding seasons had died—whereas for the remaining 40%, it reflects their fledgling production up to 2015. We used cumulative reproductive success to not exclude successful individuals that were still breeding in the last 3 years of the study (13% of the 206 individuals for which CRS does not equal LRS had been recorded for at least 15 years). We also calculated lifetime fledgling production for the subset of 161 females and 143 males that were supposedly dead in 2015, because they were not observed in 2013–2015.

Cumulative reproductive success and LRS were modelled using a GLMM with a Poisson distribution, a log-link and a Laplace approximation. As explanatory variable we added the morph of each individual. We included the first year of the breeding career of an individual as random intercept, to account for between-cohort variation and for the fact that in more recent years some individuals were still alive. Moreover, we analysed CRS by adding to the previous model the number of breeding attempts as covariate. To avoid bias in the cumulative fitness estimates due to detection rates <100%, we reran the analyses excluding individuals that were missing in the data set for more than 2 years (34 of 510 adults, 6.6%).

2.5 | Changes in morph frequencies over time

We examined temporal variation in morph frequencies for both males and females across the 20 years of the study. We used the data from all individuals for which the morph had been scored in each year ($N_{females}$ = 1,453 individual-years; N_{males} = 1,428 individual-years). Because some birds featured in multiple years, leading to pseudo-replication, we also assessed changes in morph frequencies of all individuals in their first year of breeding only ($N_{females} = 266$; $N_{males} = 244$).

We fitted a generalized linear model for each of the three morphs, where the dependent variable is the proportion of all individuals of a given morph and the independent variables are sex and year. The models were fitted with a binomial error distribution corrected for under-dispersion (i.e., using the quasi-binomial family). The robustness of the models was evaluated using a nonparametric bootstrap procedure with the function *Boot* in package *car* (Fox & Weisberg, 2011).

3 | RESULTS

3.1 | Adult apparent survival

Capture–recapture analysis showed no difference among the three morphs in the probability of resighting (p). p varied among years both in females and in males (Table S1). Based on these models, mean annual resighting rates were 0.86 (95% confidence interval, CI = 0.83–0.87) for males, and 0.87 (CI = 0.85–0.89) for females.

For both sexes, we found little support for morph-dependent survival rates, whereby the best-supported model was the one without other factors included (Table 1). For males, we found some support (delta AICc < 2) for a model that included morph, but effects of morph are at best minor (Figure 1). Based on the best-supported null model, males had similar survival (estimate = 0.90, CI = 0.88–0.91) as females (estimate = 0.88, CI = 0.86–0.90). Models that included ecological covariates reflecting yearly variation in winter severity, food availability and the investment during the previous breeding season were not better supported than the null model (Table S3; Figure S1).

3.2 | Mating patterns

Common buzzards showed weak positive assortative mating with respect to colour morph (Pearson's $r = .13$, $N = 400$ unique pairs,

TABLE 1 Results of a capture–recapture analysis for data on breeding male and female common buzzards between 1996 and 2015

Model	*N*	AICc	Delta AICc	Weight	Deviance
Males					
$\varphi(.)\ p(t)$	20	1,128.62	0.0	0.719	665.47
$\varphi(m)\ p(t)$	22	1,130.51	1.87	0.280	663.22
$\varphi(t)\ p(t)$	38	1,153.69	25.84	0.000	652.90
$\varphi(t + m)\ p(t)$	40	1,155.71	27.13	0.000	650.68
$\varphi(t \times m)\ p(t)$	76	1,222.68	94.05	0.000	638.94
Females					
$\varphi(.)\ p(t)$	20	1,942.12	0.0	0.881	1,098.37
$\varphi(m)\ p(t)$	22	1,945.14	3.01	0.119	1,097.26
$\varphi(t)\ p(t)$	38	1,956.21	14.09	0.000	1,074.85
$\varphi(t + m)\ p(t)$	40	1,959.05	16.93	0.000	1,073.46
$\varphi(t \times m)\ p(t)$	76	1,990.32	48.20	0.000	1,026.14

Note: Individuals are categorized by colour morph (dark, intermediate and light). The analysis separates between survival probabilities (φ) that can be either constant (.), morph-dependent (*m*), year-dependent (*t*), or both morph- and year-dependent ($m \times t$), and recapture probabilities (p) that are year dependent (*t*). All five possible models are displayed in decreasing order of AICc-values (fit to the data). Shown are the number of parameters (*n*), the corrected Akaike information criterion (AICc), delta AICc (the difference in AICc between the current model and the best model), the proportional support for the model (i.e. the AICc weight) and the deviance. Models are corrected for overdispersion ($\hat{C} = 1.77$ for males and $\hat{C} = 1.48$ for females).

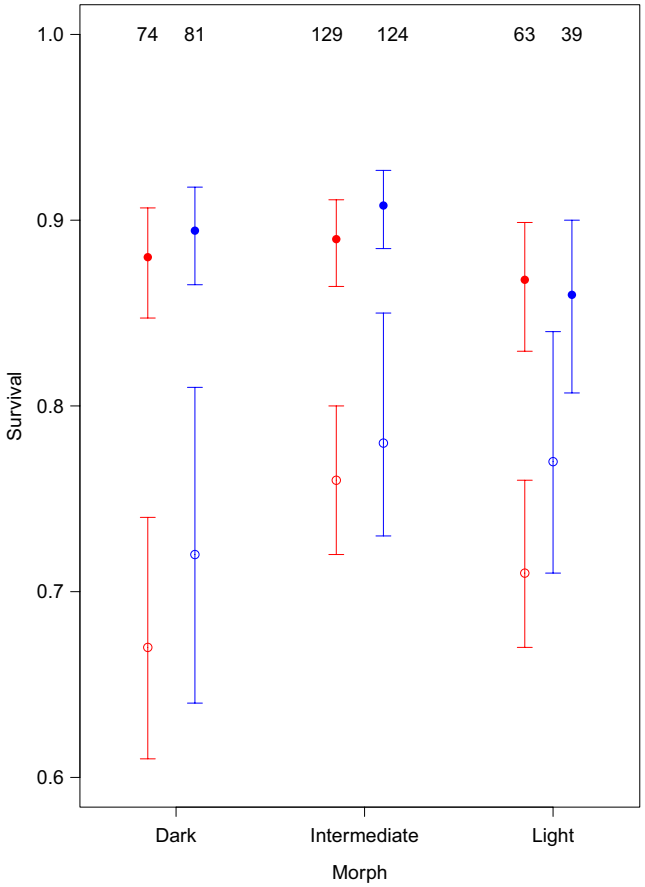

FIGURE 1 Apparent survival probabilities in relation to plumage colour morph for females (red) and males (blue). Filled symbols show the survival estimates (±95% CI) from our Dutch population, based on the capture-recapture model $\varphi(m)\ p(t)$ (see Results and Table 1). Open symbols show the average survival probabilities (±*SE*) for each morph in the German population after model averaging (data from Jonker et al., 2014: $N_{males} = 670$, $N_{females} = 669$). Numbers on top indicate sample sizes for our study

$p = .01$). The estimate of assortative mating was stronger when considering all breeding attempts ($r = .24$, $N = 1{,}566$, $p < .001$), suggesting that positively assorted pairs bred together for more years than disassortative pairs. Indeed, pair bond duration increased with the level of assortative mating ($\chi^2_2 = 16.459$, $p < .001$), where the disassortative pairs (scored as 0) had a significantly lower pair bond duration than pairs that were intermediately assorted (scored as 1) or highly assorted (scored as 2) (Figure 2; Table S4). Note that there was no difference in pair bond duration between light-light/dark-dark pairs (pooled, $N = 64$) and intermediate-intermediate pairs ($N = 108$; Table S5).

We also estimated the level of assortative mating for each year separately. This showed a significant positive assortment in thirteen out of 20 years (all $r > .18$, all $p < .05$ in 1999, 2001–2004, 2006–2007, 2009–2013, 2015; Figure S3). Interestingly, in the first 3 years of the study we found no support for positive assortative mating by morph, but levels increased and stabilized thereafter at annual correlation coefficients varying around 0.27.

3.3 | Reproduction

Neither male nor female morph explained any of the variation in the annual number of fledglings produced. However, pair assortment with respect to morph had a significant influence on reproductive success (Table S6): assortative pairs fledged more offspring (Figure 3a,b). This result was mainly driven by a difference in nest success (the probability that a brood produced at least one fledgling) among pairs with different degrees of assortment by morph (Figure 3c,d; Table S7).

Individual cumulative reproductive success was dependent on morph for both males ($\chi^2_2 = 6.04$, $p = .05$) and females ($\chi^2_2 = 13.361$, $p = .001$). In females, the intermediate morph had a significantly higher CRS than the dark morph and in males intermediates had a significantly higher CRS than both extremes (Figure 4; Table S8). When re-running the analysis on the subset of individuals that were not missing for more than 2 years from the population, we found that the intermediate morph had significantly higher CRS than the dark morph in both sexes (Table S9). Note that when we controlled for the number of breeding attempts, the effect of morph was no longer significant (Table S10). When considering LRS on the subset of individuals that were assumed dead (after not having been observed in three consecutive breeding seasons), we found no significant effect

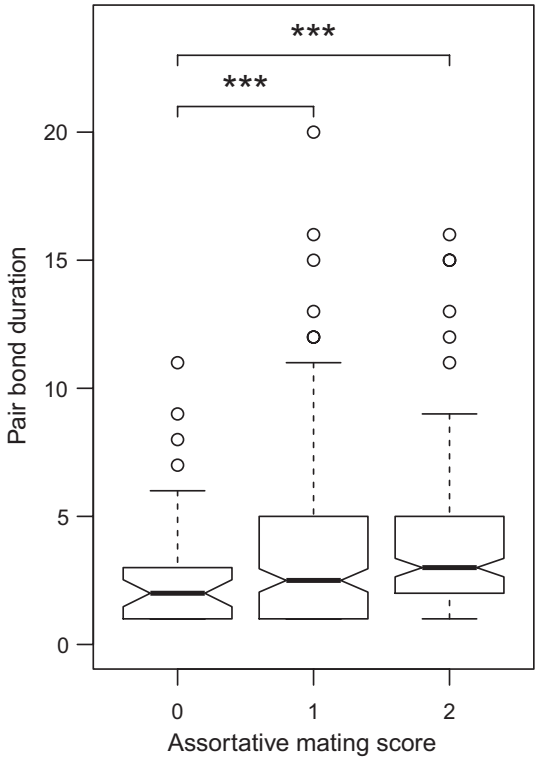

FIGURE 2 Pair bond duration in relation to the level of assortative mating based on plumage morph. Shown are boxplots of pair bond duration (the number of years a pair bred in the population) for three classes of assortment by morph: '0' for dark-light pairs ($N = 36$), '1' for intermediate-dark or intermediate-light pairs ($N = 192$), and '2' for pairs where both partners had the same morph ($N = 172$). Circles indicate statistical outliers. Asterisks indicate $p < .001$ (see Table S4). Tukey post-hoc comparisons, 0–1: $z = 3.687$, $p < .001$; 0–2: $z = 4.016$, $p < .001$; 1–2: $z = 0.676$, $p = .7$

of morph for females and a significant difference between intermediates and dark morphs for males (Table S11).

3.4 | Morph frequencies over time

Over the entire study period, 155 out of 510 individuals (30%) belonged to the dark morph, 253 (50%) belonged to the intermediate morph and 102 (20%) belonged to the light morph. Frequencies were similar for males and females: (males: 33% D, 51% I, 16% L; females: 28% D, 48% I, 24% L; numbers given in Figure 5).

Morph frequencies varied significantly among the years for both sexes, with intermediates becoming more frequent compared to the extreme morphs as the study progressed (Figure 5; Table S12; Figure S4).

4 | DISCUSSION

A previous study on common buzzards suggested that the colour polymorphism was maintained by heterosis, assuming a simple one-locus two-allele inheritance system with the intermediate morph being the heterozygote (Krüger et al., 2001). This study showed that the intermediate colour morph had the highest fitness. In agreement with Krüger et al. (2001), we found that fitness (cumulative reproductive success) differed among morphs, with the intermediate morph having highest fitness. In contrast to this earlier study, however, we found that the proportion of intermediate individuals increased over a 20-year period. This apparent evolutionary change did not just arise due to selection on individual phenotypes, but likely also from fitness benefits of assortative mating. Assortative pairs were more successful in raising offspring than disassortative pairs and assortatively paired intermediates produced a higher percentage of intermediate offspring (74%) than expected under a simple Mendelian inheritance system (50%) (Kappers et al., 2018), which could lead to a decline in frequencies of extreme phenotypes. This could lead to a positive feedback loop, for example if the extreme morphs take longer to find a suitable (assortative) mate, and ultimately to a decline in the extreme genotypes.

Because our study is a replication of the earlier studies on fitness consequences of plumage morph in common buzzards (Boerner & Krüger, 2009; Jonker et al., 2014; Krüger et al., 2001), we first discuss differences and similarities between the two studies, focussing on each of the considered fitness components. Differences between the studies could arise through differences in ecological factors between the two populations, potentially leading to different selection pressures, or through methodological differences.

In both studies, adult survival only differed minimally between the morphs, whereby the intermediate morph had slightly higher estimated annual survival (Figure 1). However, annual survival was considerably lower in the German population (Jonker et al., 2014; Figure 1). Both studies were based on large sample sizes and a long-term data set and used similar methods for analysing annual survival. However, no (morph-specific) resighting rates have been reported for the German population. The method of individual identification of the breeding buzzards differed between the studies; the German study relied mostly on visual observations and photographs, whereas we mostly used the unique banding patterns of moulted feathers. Whether this difference affects survival estimates remains unknown, but it seems unlikely that it could explain the much lower estimated survival rates in the German population. Individuals of the dark morph are probably most difficult to distinguish, but they constitute only 13% of the German population (Boerner & Krüger, 2009). Alternatively, selection pressures may be different in Germany, leading to lower local survival. This is not unlikely, given that the two populations differ in two relevant aspects. (a) The German population increased fourfold between 1989–2015 (Mueller, Chakarov, Krüger, & Hoffman, 2016), whereas the Dutch population was rather stable. (b) Eagle-owls (*Bubo bubo*) colonized the German area as predators of buzzards since 2003 (Mueller et al., 2016), but are absent in the Dutch population. In addition, a constant effort to defend against poaching in the Dutch study area could have helped in maintaining a stable number of adults holding territories. Our survival estimates are comparable to those reported for adult common buzzards from a UK population (88%–91%, Kenward et al., 2000), and they are also

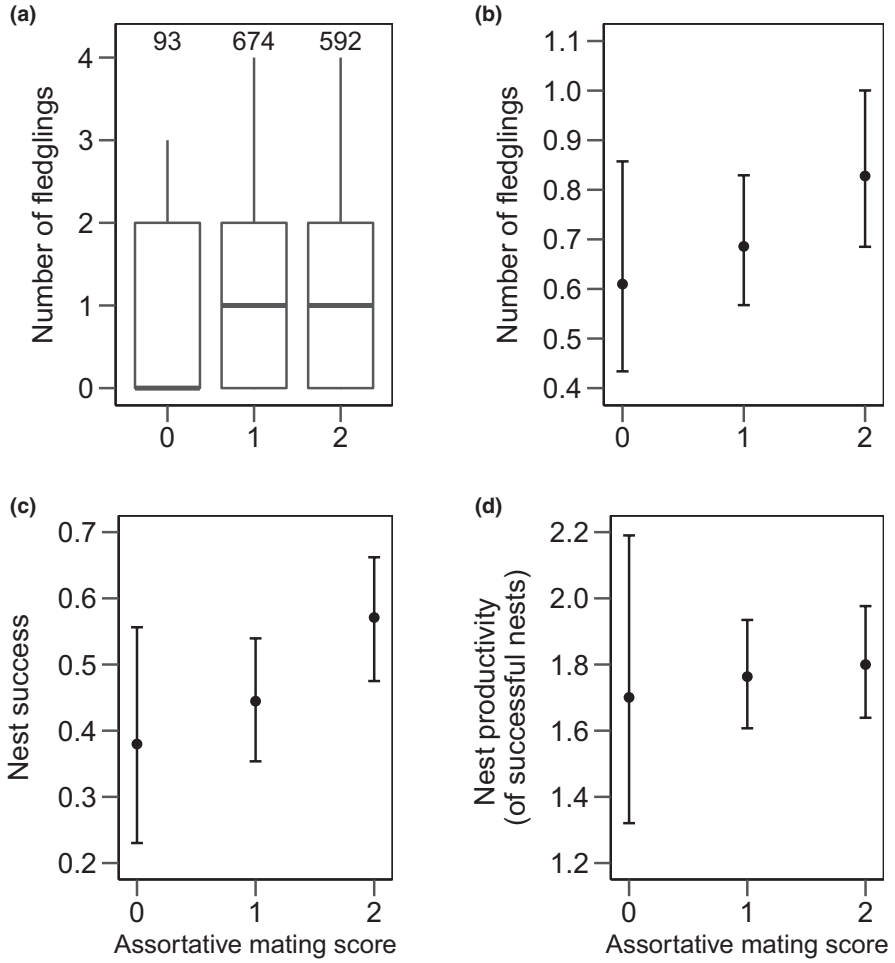

FIGURE 3 Yearly reproductive success in relation to the degree of pair assortment by morph. Three classes of assortment by morph are considered: '0' for dark-light pairs ($N = 93$), '1' for intermediate-dark or intermediate-light pairs ($N = 674$), and '2' for pairs where both partners had the same morph ($N = 592$: $N_{DxD} = 156$, $N_{IxI} = 324$, $N_{LxL} = 112$). (a) The number of fledglings per brood. Shown are boxplots of the raw data. (b) Number of fledglings per brood. Shown are means ± 95% confidence intervals from the GLMM model with female morph, male morph and disturbance as independent variables. Assortative pairs fledged more offspring (Tukey post-hoc comparisons, 0–1: $z = 0.70$, $p = .75$; 0–2: $z = 1.8$, $p = .16$; 1–2: $z = 2.33$, $p = .047$ (see Table S6). (c) Nest success, that is the probability of producing at least one fledgling (mean ± 95% CI, see Table S7). (d) The number of fledglings considering only successful nests (with at least one offspring fledged). Shown are means ± 95% CI (Table S7)

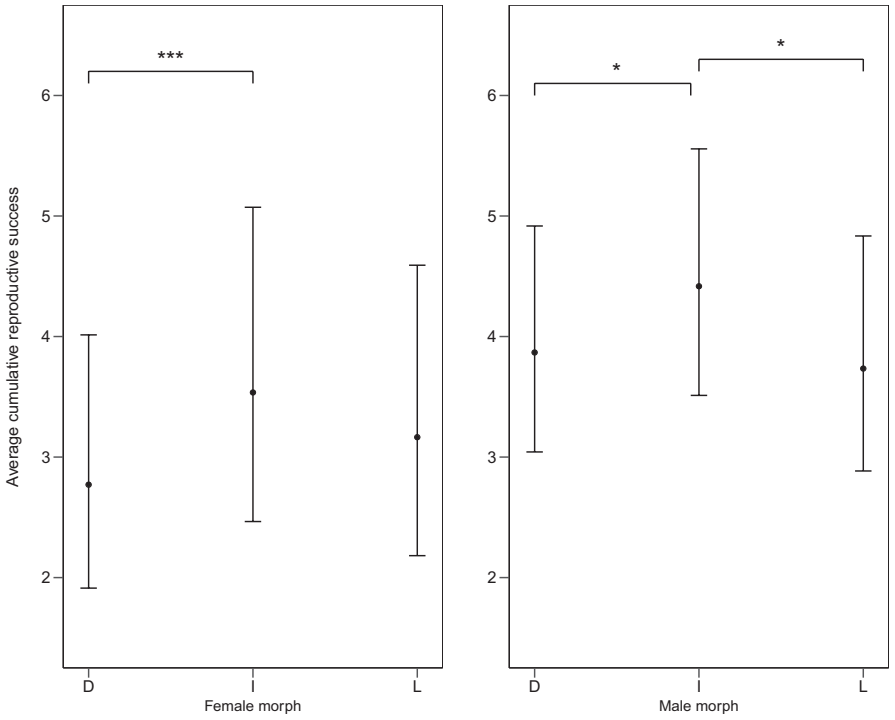

FIGURE 4 Cumulative reproductive success for females ($N = 266$) and males ($N = 244$), defined as the total number of fledglings produced during their presence in the population, in relation to individual morph (D = dark, I = intermediate, L = light). CRS was modelled using a GLMM with a Poisson distribution, a log-link and a Laplace approximation (Table S8; difference between dark and intermediate females: *$p < .001$, between dark and intermediate males: $p = 0.040$, between light and intermediate males: $p = 0.044$*). Shown are means ± 95% confidence intervals

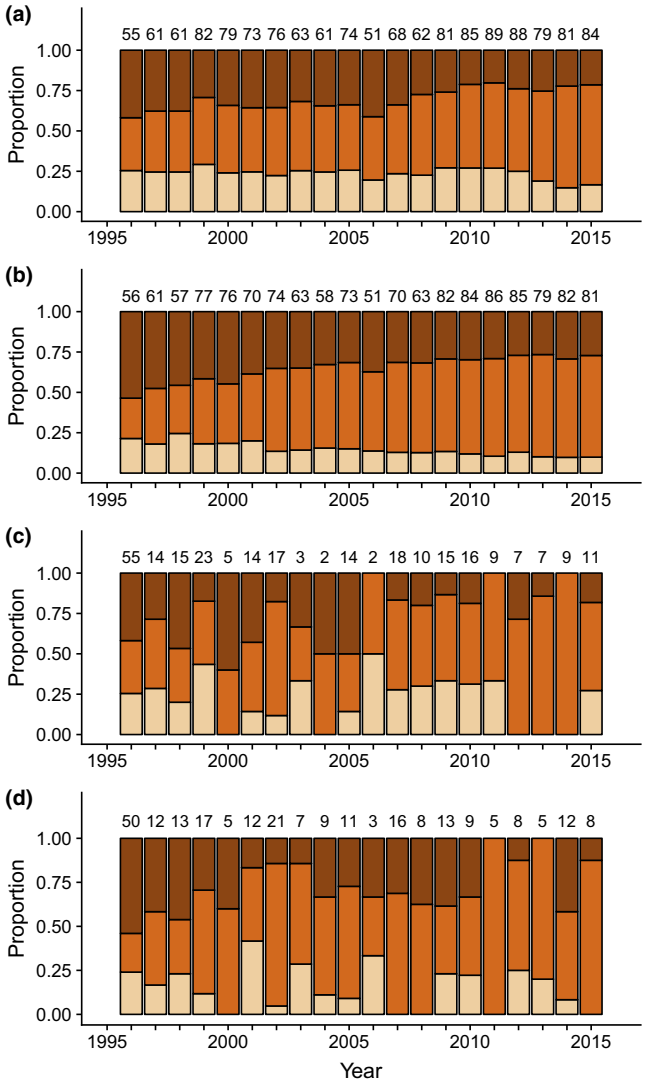

FIGURE 5 Yearly proportions of buzzards of the three plumage colour morphs between 1996 and 2015 (brown = dark morph, orange = intermediate morph, beige = light morph). (a) Data from all females in the breeding population in each year ($N_{females}$ = 1,453 individual-years). (b) Data from all males in the breeding population in each year (N_{males} = 1,428 individual-years). (c) Data from all females in their first year of breeding ($N_{females}$ = 266). (d) Data from all males in their first year of breeding (N_{males} = 244). Numbers on top indicate sample sizes

similar to survival rates observed for other medium-sized hawks (reviewed by Newton, Mcgrady, & Oli, 2016).

In the Dutch population, annual reproductive success was unrelated to morph. No comparative data have been published for the German population. However, in the Dutch population the mean number of fledglings seems about 50% lower than in the German population (Krüger, 2004). These data may not be directly comparable, because we included all pairs that held a territory, whereas Krüger (2004) only included breeding pairs.

In both populations, there is evidence for assortative mating for colour morph. Our study shows that assortative pairs were more likely to produce offspring, and that pair bonds lasted longer. In contrast, assortative mating in the German population was considered maladaptive, because under the hypothesis of simple Mendelian inheritance, light–dark pairs would produce 100% intermediate offspring with higher fitness (Krüger et al., 2001). However, this simple inheritance pattern is not consistent with the data (Kappers et al., 2018). It remains unclear why assortative pairs in the Dutch population performed better, but it might be related to behavioural compatibility or to local habitat matching. Evidence for the former comes from a study on another polymorphic raptor, the black sparrowhawk, *Accipiter melanoleucus* (Tate, Sumasgutner, Koeslag, & Amar, 2017). This study showed that neither of the two morphs had an advantage in terms of productivity or survival, but that the morph combination of adult pairs significantly influenced productivity. Mixed-morph pairs produced more offspring per year than same morph pairs, possibly due to behavioural complementarity (Tate et al., 2017). Although in this case disassortative rather than assortative pairs had higher success, the study shows that pair-level fitness advantages may play an important role in promoting and maintaining a colour polymorphism in species with biparental care.

In both populations, there is clear evidence that long-term fitness measures differ between the morphs in favour of the intermediates. However, the effect sizes were much larger in the German population, where the intermediates produced at least twice as many fledglings during their lives compared to dark or light morphs (Boerner & Krüger, 2009). In our population, intermediates had a 15% higher fitness. In the German population, the fitness differences between the morphs were due to both differences in mean life span and differences in reproductive success (Krüger et al., 2001). However, we found no significant difference in reproductive success between the morphs after controlling for the number of breeding attempts (Table S10).

Krüger et al. (2001) suggested that the higher fitness of individuals of the intermediate morph is due to (a) intermediates breeding in the highest quality territories, and (b) dark and light individuals having a lower breeding propensity. Hence, they suggested that the competitive advantage of individuals of the intermediate morph (Krüger, 2002), in combination with large variation in territory quality, resulted in the observed fitness advantage. Chakarov, Boerner, and Krüger (2008) further suggested that the success of intermediate morphs could be related to parasite resistance. The study shows that buzzard nestlings with darker plumage were more susceptible to an ectoparasite (the carnid fly, *Carnus haemapterus*), whereas nestlings infected with a blood parasite (*Leucocytozoon toddi*) showed a higher infection intensity when they had lighter plumage. This suggests that the two parasite species might exert opposite selection pressures on plumage colour of the host, such that intermediate buzzards could have an advantage (Chakarov et al., 2008). However, the results depended on offspring sex and on food availability (vole density). Thus, the role of parasites in maintaining the colour polymorphism remains unclear. The lower fitness differences between the morphs in our population could be explained if territory quality is less variable in our study area. Given that competitive abilities may differ between morphs, it would be interesting to assess morph-dependent survival

in the nest and post-fledging, and age at first breeding in relation to colour morph.

We set out to repeat a previous study and to explain the maintenance of the colour polymorphism in buzzards. We conclude that the mechanism suggested by Krüger et al. (2001) for the maintenance of this polymorphism (overdominance) seems unlikely (see also Kappers et al., 2018). Instead, our results suggest that morph frequencies have changed directionally over the past years, with an increase in the proportion of intermediates. However, we failed to identify an ecological factor to explain this apparent evolutionary change. Intriguingly, in both populations intermediates seem to have a fitness benefit, suggesting a potential for evolutionary change. Nevertheless, these populations are still highly variable for this genetically determined trait. To solve this evolutionary paradox, we need a better understanding of the ecological causes behind the fitness differences. Several key pieces of information are still missing. First, we have no knowledge about morph-specific differences in survival until breeding, and in the likelihood to obtain a breeding territory. In the dimorphic juvenile mute swans *Cygnus cygnus*, the grey morph survived better, but started breeding later in life (Conover, Reese, & Brown, 2000). Different buzzard morphs might have different early life-history strategies, countering the selection in favour of intermediate adult breeders. Second, we have little information about spatial variation in selection pressures on colour morphs (Gillespie & Turelli, 1989), and about phenotype–habitat matching (Edelaar, Siepielski, & Clobert, 2008). There is ample evidence for clines or variation in colour morphs over larger (Amar, Reynolds, Van Velden, & Briggs, 2019; Antoniazza, Burri, Fumagalli, Goudet, & Roulin, 2010) and smaller (Amar, Koeslag, Malan, Brown, & Wreford, 2014; Sordahl, 2014) spatial scales in raptors. However, there is relatively little evidence for a morph-by-habitat interaction on fitness (Dreiss et al., 2012). Our study clearly highlights that understanding the evolutionary dynamics in natural populations requires not only a long-term effort in monitoring a focal population, but also needs to include measures of fitness consequences that typically accrue outside the specific study site (dispersal and habitat choice, spatial variation in fitness parameters).

ACKNOWLEDGMENTS

We thank the landowners for permission to work on their property and those who assisted in the field. We thank Rosemarie Kentie for advice on the mark–recapture analysis and Oliver Krüger and an anonymous reviewer for constructive comments. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors have no competing interests.

AUTHORS' CONTRIBUTION

E.F.K., C.B. and B.K. designed the study. C.d.V and A.A. collected the data. S.K performed the laboratory analysis. E.F.K. and M.V. analysed the data with input from B.K. and C.B. E.F.K., C.B. and B.K. wrote the manuscript. All authors revised and approved the manuscript.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jeb.13675.

DATA AVAILABILITY STATEMENT

Data are available from the Open Science Framework at https://osf.io/s2z9r/.

ORCID

Elena Frederika Kappers https://orcid.org/0000-0002-6231-0843
Mihai Valcu https://orcid.org/0000-0002-6907-7802
Bart Kempenaers https://orcid.org/0000-0002-7505-5458

REFERENCES

Allison, A. C. (1964). Polymorphism and natural selection in human populations. *Cold Spring Harbor Symposia on Quantitative Biology*, *29*, 137–149. https://doi.org/10.1101/SQB.1964.029.01.018

Amar, A., Koeslag, A., Malan, G., Brown, M., & Wreford, E. (2014). Clinal variation in the morph ratio of Black Sparrowhawks *Accipiter melanoleucus* in South Africa and its correlation with environmental variables. *Ibis*, *156*, 627–638.

Amar, A., Reynolds, C., Van Velden, J., & Briggs, C. W. (2019). Clinal variation in morph frequency in Swainson's hawk across North America: No support for Gloger's ecogeographical rule. *Biological Journal of the Linnean Society*, *127*, 299–309. https://doi.org/10.1093/biolinnean/blz037

Antoniazza, S., Burri, R., Fumagalli, L., Goudet, J., & Roulin, A. (2010). Local adaptation maintains clinal variation in melanin-based coloration of European barn owls (*Tyto alba*). *Evolution*, *64*, 1944–1954.

Arnold, T. W. (2010). Uninformative parameters and model selection using Akaike's information criterion. *Journal of Wildlife Management*, *74*, 1175–1178. https://doi.org/10.1111/j.1937-2817.2010.tb01236.x

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bijlsma, R. (2016). Trends en broedresultaten van roofvogels in Nederland in 2015. *De Takkeling*, *24*, 5–60.

Boerner, M., & Krüger, O. (2009). Aggression and fitness differences between plumage morphs in the common buzzard (*Buteo buteo*). *Behavioral Ecology*, *20*, 180–185. https://doi.org/10.1093/beheco/arn132

Briggs, C. W., Collopy, M. W., & Woodbridge, B. (2011). Plumage polymorphism and fitness in Swainson's Hawks. *Journal of Evolutionary Biology*, *24*, 2258–2268. https://doi.org/10.1111/j.1420-9101.2011.02356.x

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York, NY: Springer.

Burnham, K. P., Anderson, D. R., White, G. C., Brownie, C., & Pollock, K. H. (1987). Design and analysis methods for fish survival experiments based on release-recapture. *American Fisheries Society Monograph*, *5*, 1–437.

Chakarov, N., Boerner, M., & Krüger, O. (2008). Fitness in common buzzards at the cross-point of opposite melanin-parasite interactions. *Functional Ecology*, *22*, 1062–1069. https://doi.org/10.1111/j.1365-2435.2008.01460.x

Conover, M. R., Reese, J. G., & Brown, A. D. (2000). Costs and benefits of subadult plumage in mute swans: Testing hypotheses for the evolution of delayed plumage maturation. *American Naturalist*, *156*, 193–200. https://doi.org/10.1086/303377

Dare, P. J. (2015). *The life of buzzards*. Dunbeath, UK: Whittles Publishing.

Dreiss, A. N., Antoniazza, S., Burri, R., Fumagalli, L., Sonnay, C., Frey, C., … Roulin, A. (2012). Local adaptation and matching habitat choice in female barn owls with respect to melanic

coloration. *Journal of Evolutionary Biology*, *25*, 103–114. https://doi.org/10.1111/j.1420-9101.2011.02407.x

Edelaar, P., Siepielski, A. M., & Clobert, J. (2008). Matching habitat choice causes directed gene flow: A neglected dimension in evolution and ecology. *Evolution*, *62*, 2462–2472. https://doi.org/10.1111/j.1558-5646.2008.00459.x

Fowlie, M. K., & Krüger, O. (2003). The evolution of plumage polymorphism in birds of prey and owls: The apostatic selection hypothesis revisited. *Journal of Evolutionary Biology*, *16*, 577–583. https://doi.org/10.1046/j.1420-9101.2003.00564.x

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*. Los Angeles, CA: SAGE Publications.

Galeotti, P., Rubolini, D., Dunn, P. O., & Fasola, M. (2003). Colour polymorphism in birds: Causes and functions. *Journal of Evolutionary Biology*, *16*, 635–646. https://doi.org/10.1046/j.1420-9101.2003.00569.x

Gillespie, J. H., & Turelli, M. (1989). Genotype-environment interactions and the maintenance of polygenic variation. *Genetics*, *121*, 129–138.

Hugall, A. F., & Stuart-Fox, D. (2012). Accelerated speciation in colour-polymorphic birds. *Nature*, *485*, 631–634. https://doi.org/10.1038/nature11050

Huxley, J. (1955). Morphism in birds. *Acta International Congress Ornithology*, *11*, 309–328.

Jones, P. D., Jonsson, T., & Wheeler, D. (1997). Extension to the North Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *International Journal of Climatology*, *17*, 1433–1450. https://doi.org/10.1002/(SICI)1097-0088(19971115)17:13<1433:AID-JOC203>3.0.CO;2-P

Jonker, R. M., Chakarov, N., & Krüger, O. (2014). Climate change and habitat heterogeneity drive a population increase in Common Buzzards *Buteo buteo* through effects on survival. *Ibis*, *156*, 97–106.

Kappers, E. F., Chakarov, N., Krüger, O., Mueller, A. K., Valcu, M., Kempenaers, B., & Both, C. (2017). Classification and temporal stability of plumage variation in Common Buzzards. *Ardea*, *105*, 125–136. https://doi.org/10.5253/arde.v105i2.a1

Kappers, E. F., de Vries, C., Alberda, A., Forstmeier, W., Both, C., & Kempenaers, B. (2018). Inheritance patterns of plumage coloration in common buzzards *Buteo buteo* do not support a one-locus two-allele model. *Biology Letters*, *14*, 20180007.

Kenward, R. E., Walls, S. S., Hodder, K. H., Pahkala, M., Freeman, S. N., & Simpson, V. R. (2000). The prevalence of non-breeders in raptor populations: Evidence from rings, radio-tags and transect surveys. *Oikos*, *91*, 271–279. https://doi.org/10.1034/j.1600-0706.2000.910207.x

Knief, U., Forstmeier, W., Pei, Y., Ihle, M., Wang, D., Martin, K., … Kempenaers, B. (2017). A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology and Evolution*, *1*, 1177–1184. https://doi.org/10.1038/s41559-017-0236-1

Krüger, O. (2002). Dissecting common buzzard lifespan and lifetime reproductive success: The relative importance of food, competition, weather, habitat and individual attributes. *Oecologia*, *133*, 474–482. https://doi.org/10.1007/s00442-002-1053-y

Krüger, O. (2004). The importance of competition, food, habitat, weather and phenotype for the reproduction of buzzard. *Bird Study*, *51*, 125–132.

Krüger, O., Lindström, J., & Amos, W. (2001). Maladaptive mate choice maintained by heterozygote advantage. *Evolution*, *55*, 1207–1214. https://doi.org/10.1111/j.0014-3820.2001.tb00640.x

Laake, J. L. (2013). *RMark: an R interface for analysis of capture-recapture data with MARK*. AFSC Process Report, 2013–01, Alaska Fisheries Science Center, NOAA.

Lewontin, R. C. (1974). *The genetic basis of evolutionary change*. New York, NY: Columbia University Press.

Mueller, A. K., Chakarov, N., Krüger, O., & Hoffman, J. I. (2016). Long-term effective population size dynamics of an intensively monitored vertebrate population. *Heredity*, *117*, 290–299. https://doi.org/10.1038/hdy.2016.67

Nakagawa, S., & Parker, T. H. (2015). Replicating research in ecology and evolution: Feasibility, incentives, and the cost-benefit conundrum. *BMC Biology*, *13*, 1–6. https://doi.org/10.1186/s12915-015-0196-3

Newton, I., Mcgrady, M. J., & Oli, M. K. (2016). A review of survival estimates for raptors and owls. *Ibis*, *158*, 227–248. https://doi.org/10.1111/ibi.12355

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Roulin, A. (2004). The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. *Biological Reviews*, *79*, 815–848. https://doi.org/10.1017/S1464793104006487

Sinervo, B., & Calsbeek, R. (2006). The developmental, physiological, neural, and genetical causes and consequences of frequency-dependent selection in the wild. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 581–610. https://doi.org/10.1146/annurev.ecolsys.37.091305.110128

Smith, J. M. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.

Sordahl, T. A. (2014). Distribution of color-morphs of the Eastern Screech-Owl in Iowa. *The Wilson Journal of Ornithology*, *126*, 321–332. https://doi.org/10.1676/13-140.1

Tate, G., Sumasgutner, P., Koeslag, A., & Amar, A. (2017). Pair complementarity influences reproductive output in the polymorphic black sparrowhawk *Accipiter melanoleucus*. *Journal of Avian Biology*, *48*, 387–398.

White, G. C., & Burnham, K. P. (1999). Program MARK: Survival estimation from populations of marked animals. *Bird Study*, *46*, S120–S139. https://doi.org/10.1080/00063659909477239

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Kappers EF, de Vries C, Alberda A, et al. Morph-dependent fitness and directional change of morph frequencies over time in a Dutch population of Common buzzards *Buteo buteo*. *J Evol Biol*. 2020;33:1306–1315. https://doi.org/10.1111/jeb.13675